



Efficacy and safety of bimekizumab in patients with moderate-to-severe hidradenitis suppurativa (BE HEARD I and BE HEARD II): two 48-week, randomised, double-blind, placebo-controlled, multicentre phase 3 trials

Alexa B Kimball*, Gregor B E Jemec*, Christopher J Sayed, Joslyn S Kirby, Errol Prens, John R Ingram, Amit Garg, Alice B Gottlieb, Jacek C Szepietowski, Falk G Bechara, Evangelos J Giamarellou-Bourboulis, Hideki Fujita, Robert Roller, Paulatsya Joshi, Pratiksha Dokhe, Edward Muller, Luke Peterson, Cynthia Madden, Muhammad Bari, Christos C Zouboulis

Summary

Background Patients with hidradenitis suppurativa have substantial unmet clinical needs and scarce therapeutic options. We aimed to assess the efficacy and safety of bimekizumab, a monoclonal IgG1 antibody that selectively inhibits interleukin (IL)-17F and IL-17A, in patients with moderate-to-severe hidradenitis suppurativa.

Methods BE HEARD I and II were two identically designed, 48-week randomised, double-blind, placebo-controlled, multicentre phase 3 trials. Patients aged 18 years or older with moderate-to-severe hidradenitis suppurativa were randomly assigned 2:2:2:1 using interactive response technology (stratified by worst Hurley Stage at baseline and baseline systemic antibiotic use) to receive subcutaneous bimekizumab 320 mg every 2 weeks; bimekizumab 320 mg every 2 weeks to week 16, then every 4 weeks to week 48; bimekizumab 320 mg every 4 weeks to week 48; or placebo to week 16, then bimekizumab 320 mg every 2 weeks. The primary outcome was an hidradenitis suppurativa clinical response of at least 50%, defined as a reduction in total abscess and inflammatory nodule count of at least 50% from baseline with no increase from baseline in abscess or draining tunnel count (HiSCR50) at week 16. Efficacy analyses included all randomly assigned study patients (intention-to-treat population). Safety analyses included all patients who received at least one full or partial dose of study treatment in the safety set, and of bimekizumab in the active-medication set. These trials are registered at ClinicalTrials.gov, NCT04242446 and NCT04242498, and both are completed.

Findings Patients for BE HEARD I were recruited from Feb 19, 2020, to Oct 27, 2021, and 505 patients were enrolled and randomly assigned. Patients for BE HEARD II were recruited from March 2, 2020, to July 28, 2021, and 509 patients were enrolled and randomly assigned. The primary outcome at week 16 was met in the group who received bimekizumab every 2 weeks using modified non-responder imputation; higher responder rates were observed with bimekizumab versus placebo in both trials: 138 (48%) of 289 patients versus 21 (29%) of 72 patients in BE HEARD I (odds ratio [OR] 2·23 [97·5% CI 1·16–4·31]; $p=0\cdot0060$) and 151 (52%) of 291 patients versus 24 (32%) of 74 patients in BE HEARD II (2·29 [1·22–4·29]; $p=0\cdot0032$). In BE HEARD II, HiSCR50 was also met in the group who were administered bimekizumab every 4 weeks (77 [54%] of 144 vs 24 [32%] of 74 with placebo; 2·42 [1·22–4·80]; $p=0\cdot0038$). Responses were maintained or increased to week 48. Serious treatment-emergent adverse events were reported in 40 (8%) patients in BE HEARD I and in 24 (5%) patients in BE HEARD II treated with bimekizumab over 48 weeks. The most frequently reported treatment-emergent adverse events to week 48 were hidradenitis in both trials, in addition to coronavirus infection and diarrhoea in BE HEARD I, and oral candidiasis and headache in BE HEARD II. One death was reported across the two trials, and was due to congestive heart failure in a patient with substantial cardiovascular history treated with bimekizumab every 2 weeks in BE HEARD I (considered unrelated to bimekizumab treatment by the investigator). No new safety signals were observed.

Interpretation Bimekizumab was well tolerated by patients with hidradenitis suppurativa and produced rapid and deep clinically meaningful responses that were maintained up to 48 weeks. Data from these two trials support the use of bimekizumab for the treatment of patients with moderate-to-severe hidradenitis suppurativa.

Funding UCB Pharma.

Copyright © 2024 Elsevier Ltd. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Introduction

Hidradenitis suppurativa is a chronic, relapsing inflammatory skin disease associated with substantial

comorbidities and a detrimental effect on the quality of life of patients.^{1–3} Painful inflammatory nodules, abscesses, and draining tunnels in folding areas of the

Lancet 2024; 403: 2504–19

Published Online

May 22, 2024

[https://doi.org/10.1016/S0140-6736\(24\)00101-6](https://doi.org/10.1016/S0140-6736(24)00101-6)

See [Comment](#) page 2457

*Joint first authors

Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

(Prof A B Kimball MD);

Department of Dermatology, Zealand University Hospital, Roskilde, Denmark

(Prof G B E Jemec DMSc);

Department of Clinical Medicine, Faculty of Health and Medical Science, University of Copenhagen, Copenhagen, Denmark (Prof G B E Jemec);

Department of Dermatology, University of North Carolina School of Medicine, Chapel Hill, NC, USA (C J Sayed MD);

European Hidradenitis Suppurativa Foundation,

Dessau, Germany

(Prof G B E Jemec, C J Sayed,

Prof E Prens PhD, J R Ingram DM,

Prof J C Szepietowski PhD,

Prof F G Bechara MD,

Prof E J Giamarellou-Bourboulis MD,

Prof C C Zouboulis PhD);

Department of Dermatology,

Penn State University, Hershey,

PA, USA (Prof J S Kirby MD);

Department of Dermatology,

Erasmus University Medical

Centre Rotterdam, Rotterdam,

Netherlands (Prof E Prens);

Department of Dermatology

and Academic Wound Healing,

Division of Infection and

Immunity, Cardiff University,

Cardiff, UK (J R Ingram);

Northwell Health, New Hyde

Park, NY, USA (Prof A Garg PhD);

Department of Dermatology,

Icahn School of Medicine at

Mount Sinai, New York,

NY, USA (Prof A B Gottlieb MD);

Department of Dermatology,

Venereology, and Allergy,

Research in context

Evidence before this study

Hidradenitis suppurativa is a chronic, systemic, relapsing inflammatory skin disease associated with disability and comorbidities, a detrimental impact on patients' quality of life, and increased risk of depression and suicidality. We searched PubMed with the term "hidradenitis suppurativa" and screened by title to identify industry-sponsored clinical trials and systematic literature reviews of biological agents in patients with hidradenitis suppurativa. Manuscripts published between Jan 1, 2006, and Dec 31, 2021, were extracted. Webpages of pharmaceutical companies with new molecules for hidradenitis suppurativa in their pipelines were also screened. Patients with hidradenitis suppurativa face substantial unmet clinical need; however, the only biological therapies currently approved are the TNF α inhibitor adalimumab and, since this literature review was performed, the interleukin (IL)-17A inhibitor secukinumab. As important pathogenic drivers, both IL-17A and IL-17F are potential therapeutic targets in hidradenitis suppurativa. Bimekizumab is a monoclonal IgG1 antibody that selectively inhibits IL-17F and IL-17A. In a phase 2 trial, bimekizumab showed clinically meaningful and consistent improvements versus placebo, including in stringent outcome measures, and was well tolerated. On the basis of these results, phase 3 studies were initiated.

Added value of this study

Bimekizumab is a first-in-class biological agent for inhibition of IL-17F and IL-17A, and has demonstrated clinically meaningful improvements in physician-assessed and patient-reported outcome measures to 48 weeks of treatment. The BE HEARD I

and II trials showed that patients with hidradenitis suppurativa treated with bimekizumab reached rapid and maintained improvements in the signs and symptoms of disease, including the primary outcome of HiSCR50, versus those who received placebo. Rapid improvements were also observed in patients who switched from placebo to bimekizumab treatment. The trials were the first phase 3 studies in hidradenitis suppurativa to report the more stringent HiSCR75 and HiSCR90 outcomes at a longer term (to week 48), and demonstrated deep and maintained clinical response with bimekizumab. The safety profile of bimekizumab in BE HEARD I and II was consistent with other bimekizumab indications, and with the selective IL-17A inhibitor, secukinumab, which is approved for use in hidradenitis suppurativa; no new safety signals were identified.

Implications of all the available evidence

The BE HEARD I and II trials are the first phase 3 trials to assess the effects of inhibition of IL-17F and IL-17A in patients with hidradenitis suppurativa. The outcomes of these trials support the hypothesised roles of both IL-17F and IL-17A in the pathogenesis of the disease, and support bimekizumab as a promising new therapeutic option for patients with moderate-to-severe hidradenitis suppurativa. Given the heterogeneity and complexity of hidradenitis suppurativa, future research should include studies targeting optimal pharmacological, surgical, and adjuvant therapies to optimise treatment goals meaningful to patients. Real-world evidence studies and network meta-analyses could also help to inform future clinician decision making in the management of moderate-to-severe hidradenitis suppurativa.

skin are the defining manifestations of hidradenitis suppurativa, which affects around 0·4–1·0% of the global population and disables as many as 14·5% of affected patients.^{1,4,5} Poor disease recognition results in substantially delayed diagnosis and intervention. Hidradenitis suppurativa affects not only skin-related quality of life, but also physical and mental health. Depression and anxiety affect up to 42·9% of patients, in whom the incidence of completed suicide is higher than that in the general population.^{2,6} Unemployment rates are high, and absenteeism is reported by around half of patients with jobs.^{5,7}

The only biological therapies approved for treatment of moderate-to-severe hidradenitis suppurativa are the tumour necrosis factor α (TNF- α) inhibitor adalimumab and the interleukin (IL)-17A inhibitor secukinumab.^{8–10} A multinational study⁵ reported that nearly half of clinicians and patients expressed dissatisfaction with existing medical interventions, showcasing the unmet need for novel effective therapies that provide rapid and sustained responses.

The pathophysiology of hidradenitis suppurativa is complex and involves immune activation with progression to chronic inflammation.^{3,11} IL-17A and IL-17F are closely

related proinflammatory cytokines that synergise with other proinflammatory cytokines to drive neutrophil influx into the lesions caused by inflammation, a key hallmark of hidradenitis suppurativa.¹² Distinct IL-17-secreting cells can be found in lesional hidradenitis suppurativa tissues, including cells that produce only IL-17A and IL-17F.¹³ Although IL-17A and IL-17F have overlapping biology in humans, with both isoforms upregulated in hidradenitis suppurativa, they are regulated differently via duration of stimulation. STAT5-inducing cytokines (eg, IL-2, IL-7, and IL-15) and IL-1 β (which is upregulated in hidradenitis suppurativa) preferentially drive IL-17-secreting cells to produce IL-17F, which might explain why IL-17F is more highly upregulated than IL-17A in hidradenitis suppurativa lesional tissue.^{13,14}

Bimekizumab is a monoclonal IgG1 antibody that selectively inhibits IL-17F in addition to IL-17A, which, compared with inhibitors of IL-17A alone, results in the additional inhibition of the IL-17F–F isomer.^{15,16} Dual inhibition of IL-17F and IL-17A in human in-vitro models of hidradenitis suppurativa with bimekizumab has been shown to more effectively suppress the production of proinflammatory cytokines than inhibition of either isoform alone and, at phase 3, bimekizumab showed

Wrocław Medical University, Wrocław, Poland (Prof J C Szepletowski); Department of Dermatology, Venereology, and Allergology, St Josef-Hospital, Ruhr-University, Bochum, Germany (Prof F G Bechara); Fourth Department of Internal Medicine, National and Kapodistrian University of Athens, Medical School, Athens, Greece (Prof E J Giamarellos-Bourboulis); Division of Cutaneous Science, Department of Dermatology, Nihon University School of Medicine, Tokyo, Japan (H Fujita PhD); UCB Pharma, Morrisville, NC, USA (R Roller PharmD, L Peterson MS, C Madden MD); UCB Pharma, Slough, UK (P Joshi MD, P Dokhe MD, E Muller MSc, M Bari MD); Departments of Dermatology, Venereology, Allergology, and Immunology, Staedtisches Klinikum Dessau, Brandenburg Medical School Theodor Fontane and Faculty of Health Sciences Brandenburg, Dessau, Germany (Prof C C Zouboulis PhD)

Correspondence to: Prof Alexa B Kimball, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02115, USA clears@bidmc.harvard.edu

superior efficacy over the selective IL-17A inhibitor secukinumab in patients with moderate-to-severe plaque psoriasis.^{16,17} Bimekizumab has also shown efficacy in patients with hidradenitis suppurativa; in a phase 2 study, clinically meaningful and consistent improvement in hidradenitis suppurativa clinical response (HiSCR) versus placebo was shown.¹⁸ On the basis of these findings, the efficacy and safety of bimekizumab were assessed in two phase 3 clinical trials in patients with moderate-to-severe hidradenitis suppurativa. As per other phase 3 programmes,^{9,10} we conducted two independent, confirmatory trials across separate centres. Side-by-side results from each trial are presented here through 48 weeks of treatment. The primary objective of these trials was to evaluate the efficacy of bimekizumab in patients with moderate-to-severe hidradenitis suppurativa.

Methods

Study design and patients

BE HEARD I and II were randomised, double-blind, placebo-controlled, multicentre phase 3 trials done in 86 sites (BE HEARD I) and 90 sites (BE HEARD II) across Europe, the USA, Canada, Asia, and Australia (appendix 1, pp 3–9). All methodological details described here apply to both BE HEARD I and BE HEARD II, unless otherwise specified. Adults (aged 18 years or older) with moderate-to-severe hidradenitis suppurativa were eligible. Moderate-to-severe disease was defined as presence of at least five inflammatory lesions (abscesses, inflammatory nodules, or both) affecting at least two distinct anatomical areas, one of which was at least Hurley stage II or III (at both screening and baseline visits), evidenced by clinical history and physical examination, and diagnosed at least 6 months before the baseline visit. Eligible patients also had a documented history of inadequate response as assessed by a physician to systemic antibiotics for hidradenitis suppurativa (eg, tetracyclines, clindamycin, and rifampicin) at screening. Patients using a stable dose (pro re nata use not accepted) of doxycycline, minocycline, or an equivalent systemic tetracycline for 28 days before baseline were allowed to continue antibiotics and enrol in the studies alongside patients not on antibiotics. Patients were excluded if they had more than 20 draining tunnels at baseline, had another active skin disease or condition that could interfere with hidradenitis suppurativa assessment, had received TNF- α inhibitors within 12 weeks or IL-17 biological response modifier therapy within 6 months of baseline, topical therapy within 14 days of baseline, or were on systemic therapy for hidradenitis suppurativa. Full eligibility and exclusion criteria are detailed in appendix 1 (pp 10–13).

The study protocol, amendments, and patient informed consent were reviewed by a national, regional, or independent ethics committee or institutional review board. This study was done in accordance with the

current version of the applicable regulatory and International Conference of Harmonisation Good Clinical Practice requirements, the ethical principles of the Declaration of Helsinki, and local laws of involved countries. Ethical approval was obtained from relevant institutional review boards at participating sites. All patients provided written informed consent in accordance with local requirements.

Randomisation and masking

Patients were randomly assigned (2:2:2:1) using interactive response technology to receive bimekizumab 320 mg every 2 weeks to week 48; bimekizumab 320 mg every 2 weeks to week 16, followed by every 4 weeks to week 48; bimekizumab 320 mg every 4 weeks to week 48; or placebo to week 16 followed by bimekizumab 320 mg every 2 weeks to week 48, on the basis of a predetermined production randomisation and packaging schedule provided by the funder. Randomisation was stratified by worst Hurley Stage at baseline (II or III) and baseline systemic antibiotic use (yes or no). For analyses that only include the initial treatment period, the groups given bimekizumab 320 mg every 2 weeks were pooled, per pre-specified analysis. For any analyses that include the maintenance treatment period, the four dose regimens are presented separately. To maintain double blinding, all patients received injections every 2 weeks to week 46. Throughout the study, patients, investigators, and the sponsor remained masked to treatment assignment except for staff needed for study drug administration and reconciliation (further details in appendix 2 pp 40, 164).

Procedures

Bimekizumab was supplied in a 1 mL prefilled syringe at a concentration of 160 mg/mL, and placebo was supplied as a 1 mL prefilled syringe of 0.9% sodium chloride aqueous solution for injection. Study treatments were administered in the clinic as two subcutaneous injections by study personnel who were not masked to treatment allocation. Because of differences in presentation between bimekizumab and placebo treatments, special precautions were taken at each study site. These precautions included having masked and unmasked personnel. Unmasked personnel administered the injections according to the site-specific blinding plan. The unmasked study personnel were also responsible for recording the administration information on source documents. All patients received injections of either bimekizumab or placebo every 2 weeks according to their group schedule. At week 16 (start of the 32-week maintenance treatment period), patients randomly assigned to bimekizumab 320 mg every 2 weeks or every 4 weeks for 48 weeks continued receiving their assigned doses. Patients randomly assigned to bimekizumab 320 mg every 2 weeks followed by 320 mg every 4 weeks started receiving injections every 4 weeks. Patients who originally received placebo received bimekizumab

See Online for appendix 1

See Online for appendix 2

320 mg every 2 weeks for the remaining duration of the 48 weeks (appendix 1 p 25).

Lesion counts and Dermatology Life Quality Index (DLQI, which assesses general skin-related quality of life) assessments were made at baseline (week 0), at weeks 4, 8, 12, 16, 20, 32, 36, and 48, and in the event of a premature end of treatment visit (appendix 2 pp 15–20, 139–144). Skin pain was assessed in two prespecified secondary outcomes using the newly developed, validated hidradenitis suppurativa symptom daily diary (HSSDD); absolute change from baseline in skin pain was assessed by the worst pain item of the HSSDD and pain response was defined as a decrease from baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change. The HSSDD is a five-item, hidradenitis suppurativa-specific patient-reported outcome tool developed by UCB Pharma in line with US Food and Drug Administration guidance, that assesses patients' perception of the core symptoms of hidradenitis suppurativa (worst skin pain, average skin pain, smell or odour, itch at its worst, and drainage or oozing from hidradenitis suppurativa lesions) experienced in the past 24 h.¹⁹ Each item is rated by the patient on an 11-point numerical rating scale, from 0 (no symptoms) to 10 (symptoms as bad as you can imagine). For each item, the HSSDD score is derived from the weekly averages of the daily scores from a given week. Higher scores indicate a higher symptom burden. A weekly HSSDD score for each item was only calculated if at least four non-missing daily values were available, otherwise the HSSDD score for the given item was reported as missing. Safety was assessed at baseline and each study visit, which occurred every two weeks. The safety follow-up visit was done 20 weeks after final dose of study treatment in patients who did not enter the subsequent open-label extension study (NCT04901195), or who prematurely withdrew. Rescue treatment for hidradenitis suppurativa, including but not limited to systemic antibiotics, intralesional injections of triamcinolone, and incision and drainage, were permitted if required as judged by the investigator (appendix 2 pp 45, 169).

Outcomes

The primary outcome of each trial was HiSCR50, assessed independently at week 16, and defined as a reduction in total abscess and inflammatory nodule count of at least 50% from baseline, with no increase from baseline in abscess and inflammatory nodule or draining tunnel count.²⁰ Key secondary outcomes in ranked testing order were as follows: attainment of a HiSCR75 response (ie, a $\geq 75\%$ reduction from baseline in total abscess and inflammatory nodule count, with no increase from baseline in abscess and inflammatory nodule or draining tunnel count) at week 16; at least one occurrence of flare (defined as a $\geq 25\%$ increase in abscess and inflammatory nodule count with an increase of at

least two abscess and inflammatory nodules relative to baseline) by week 16 (BE HEARD II only); absolute change from baseline in DLQI score at week 16 (minimum clinically important difference defined as a four-point reduction in DLQI total score);²¹ absolute change from baseline in skin pain score at week 16, assessed by the worst skin pain item (11-point numeric rating scale on the HSSDD) and hidradenitis suppurativa skin pain response at week 16 on the basis of a threshold for clinically meaningful change (defined as a within-patient reduction in HSSDD weekly worst skin pain score of at least 3 points from baseline among patients with a baseline score of at least 3). Secondary outcomes were assessed at the individual trial level.

Additional prespecified exploratory outcomes evaluated the long-term efficacy of bimekizumab measured by HiSCR50, HiSCR75, HiSCR90 (a reduction in HiSCR of at least 90% from baseline), and HiSCR100 (a reduction in HiSCR of 100% from baseline), change from baseline in abscess and inflammatory nodule count, and change from baseline in draining tunnel count. Other prespecified exploratory outcomes are listed in appendix 1 (pp 15–16).

Treatment-emergent adverse events were reported for all study groups from weeks 0–16 and for all patients who received bimekizumab treatment from baseline to week 48. Occurrence of treatment-emergent adverse events, including serious treatment-emergent adverse events and those leading to treatment discontinuation were evaluated. Incidence of adverse events and serious adverse events throughout the trials were coded using Medical Dictionary for Regulatory Activities version 19.0.

Prespecified safety events of interest included infection (serious, opportunistic, fungal, and tuberculosis), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behaviour, major adverse cardiovascular events, hepatic events, malignancies, and inflammatory bowel disease. Inflammatory bowel disease, liver function test elevations, suicidal ideation and behaviour, and major adverse cardiovascular events were adjudicated by independent external committees. An independent data monitoring committee periodically assessed safety data and provided recommendations on study conduct and safety data analyses during studies; in addition, unblinded efficacy data analysis was provided when approximately 66% of study participants reached the primary efficacy endpoint at week 16.

Statistical analyses

The primary objective was to evaluate the efficacy of bimekizumab compared with placebo in patients with moderate-to-severe hidradenitis suppurativa, by assessing HiSCR50 response at week 16. Study power was calculated for the primary outcome on the assumption that, in each trial, responder rates for HiSCR50 at week 16 would be 60% for bimekizumab every 2 weeks, 50% for bimekizumab every 4 weeks,

and 25% for placebo. The assumed responder rates for these calculations accounted for a dropout rate of approximately 10%. Using these assumptions, the power provided by a sample size of 490 patients per trial (280 patients for bimekizumab every 2 weeks, 140 patients for bimekizumab every 4 weeks, and 70 patients for placebo) to demonstrate statistical superiority of bimekizumab relative to placebo at a two-sided significance level of 0.025 was 99% for bimekizumab every 2 weeks and 90% for bimekizumab every 4 weeks. Both trials were independently powered to test primary and ranked secondary outcomes.

Efficacy analyses, including for the primary outcome, included all randomly assigned patients (intention-to-treat population). Safety analyses included all patients who received at least one full or partial dose of study treatment in the safety set, and of bimekizumab in the active medication set. Multiplicity and type I error were controlled for primary and secondary efficacy outcomes using a fixed-sequence closed testing procedure under a parallel gatekeeping framework. Evaluation of statistical significance for each outcome in the sequence was dependent on the previous comparison reaching statistical significance with a two-sided α level of 0.025, for which the two bimekizumab dose regimens were tested independently versus placebo (the testing hierarchies for both studies are presented in appendix 1 pp 26–27).

For the primary and secondary outcome analyses of HiSCR50, HiSCR75, HSSDD worst skin pain response, and flare (which was a secondary outcome in BE HEARD II only, and classified as other outcome in BE HEARD I) at week 16, odds ratios (OR; including 97.5% CI) versus placebo and p values (from Wald test) were obtained from a logistic regression model (appendix 1 p 14). For change from baseline in the secondary outcomes of DLQI score and HSSDD worst skin pain response at week 16, least-squares-mean differences (including 97.5% CI) versus placebo and p values were based on an analysis of covariance model (appendix 1 p 14).

For the primary analysis of binary outcomes specified in the statistical testing hierarchy, we used a modified non-responder imputation (mNRI) whereby patients who took any systemic antibiotic (new or increased dose for any indication) or who discontinued study treatment because of an adverse event or absence of efficacy were treated as non-responders at all subsequent visits, and other missing data were imputed using a multiple imputation (appendix 1 p 14) model (mNRI [all-antibiotics]; appendix 1 p 28). For the primary analysis of continuous outcomes specified in the statistical testing hierarchy, we used multiple imputation (all-antibiotics) whereby patients who took any systemic antibiotic (new or increased dose for any indication) or who discontinued study treatment because of an adverse event or absence of efficacy were treated as missing and subsequently

imputed using multiple imputation, and other missing data were also imputed using multiple imputation. For HiSCR50, HiSCR75, HiSCR90, and HiSCR100 responses over time, in addition to other binary exploratory outcomes, an alternative post-hoc mNRI was used, in which only patients who took systemic antibiotics identified as rescue medication for hidradenitis suppurativa by the principal investigator or who discontinued because of an adverse event or absence of efficacy were treated as non-responders at all subsequent visits, and other missing data were imputed using multiple imputation (mNRI [hidradenitis suppurativa-antibiotics]). For continuous exploratory outcomes, patients who took systemic antibiotics identified as rescue medication for hidradenitis suppurativa by the principal investigator or who discontinued study treatment because of an adverse event or absence of efficacy were treated as missing and subsequently imputed using multiple imputation (hidradenitis suppurativa-antibiotics). For all exploratory outcomes, other missing data were imputed via multiple imputation. For multiply imputed binary variables, the rounded average number of patients with response based on 100 imputations is reported. Observed case analyses are also presented whereby only data for patients on treatment were considered and missing data were not imputed.

All analyses were done using SAS version 9.4. Both BE HEARD I (NCT04242446) and BE HEARD II (NCT04242498) trials are registered with ClinicalTrials.gov and are completed.

Role of the funding source

UCB Pharma contributed to the study design, participated in data collection, completed the data analysis, and participated in data interpretation. UCB Pharma participated in writing, revision, and approval of the manuscript. A medical writing agency, employed by UCB Pharma, assisted with manuscript preparation under the authors' direction.

Results

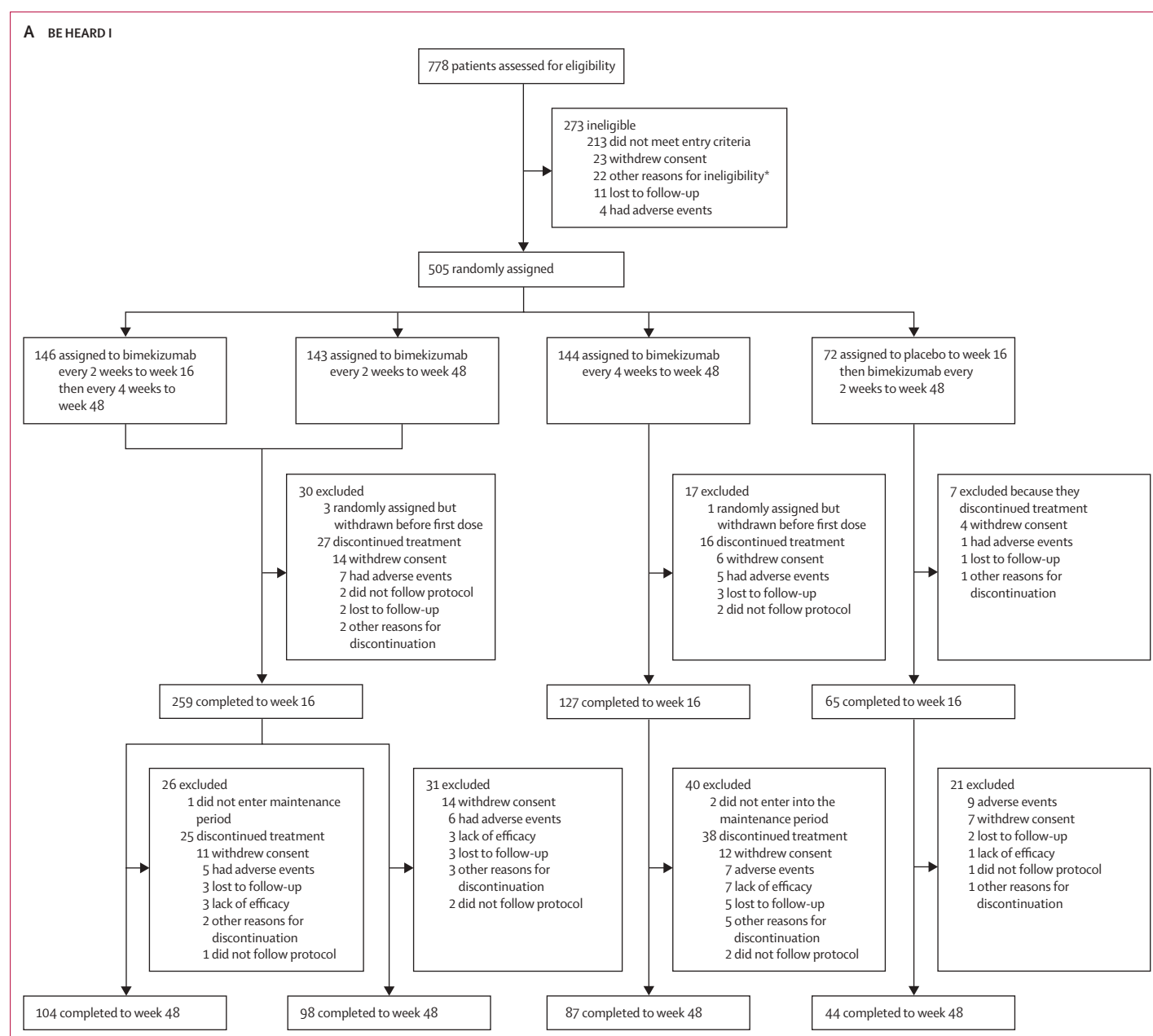
Patients for BE HEARD I were recruited from Feb 19, 2020 and Oct 27, 2021, patients for BE HEARD II were recruited from March 2, 2020 and July 28, 2021. For both trials, because of the COVID-19 pandemic, all but one patient (enrolled in March, 2020) were enrolled beginning in June, 2020. Of the 778 patients screened in BE HEARD I, 505 were randomly assigned to receive bimekizumab every 2 weeks (n=289), bimekizumab every 4 weeks (n=144), or placebo (n=72; figure 1A). Of the 726 patients screened in BE HEARD II, 509 were randomly assigned to receive bimekizumab every 2 weeks (n=291), bimekizumab every 4 weeks (n=144), or placebo (n=74; figure 1B). Baseline demographics were generally representative of patients with moderate-to-severe hidradenitis suppurativa (table 1). Overall, 451 (89%) of 505 patients in BE HEARD I and 464 (91%)

of 509 patients in BE HEARD II completed treatment to week 16, and 333 (74%) of 448 patients and 387 (84%) of 463 patients completed treatment to week 48 (figure 1; appendix 1 pp 17–18).

In BE HEARD I, the median time on study medication was 334 days (IQR 252–336) in patients who received bimekizumab every 2 weeks; 335 days (272–336) in patients who received bimekizumab every 2 weeks followed by every 4 weeks; 333 days (194–336) in patients who received bimekizumab every 4 weeks; and 336 days (280–336) in patients who received placebo

followed by bimekizumab every 2 weeks. In BE HEARD II, the median time on study medication was 334 days (314–336) in patients who received bimekizumab every 2 weeks; 334 days (267–336) in patients who received bimekizumab every 2 weeks followed by every 4 weeks; 335 days (306–336) in patients who received bimekizumab every 4 weeks; and 336 days (334–336) in patients who received placebo followed by bimekizumab every 2 weeks.

The primary outcome of HiSCR50 was met in the group of patients who received bimekizumab every 2 weeks in both trials, with higher rates of HiSCR50



(Figure 1 continues on next page)

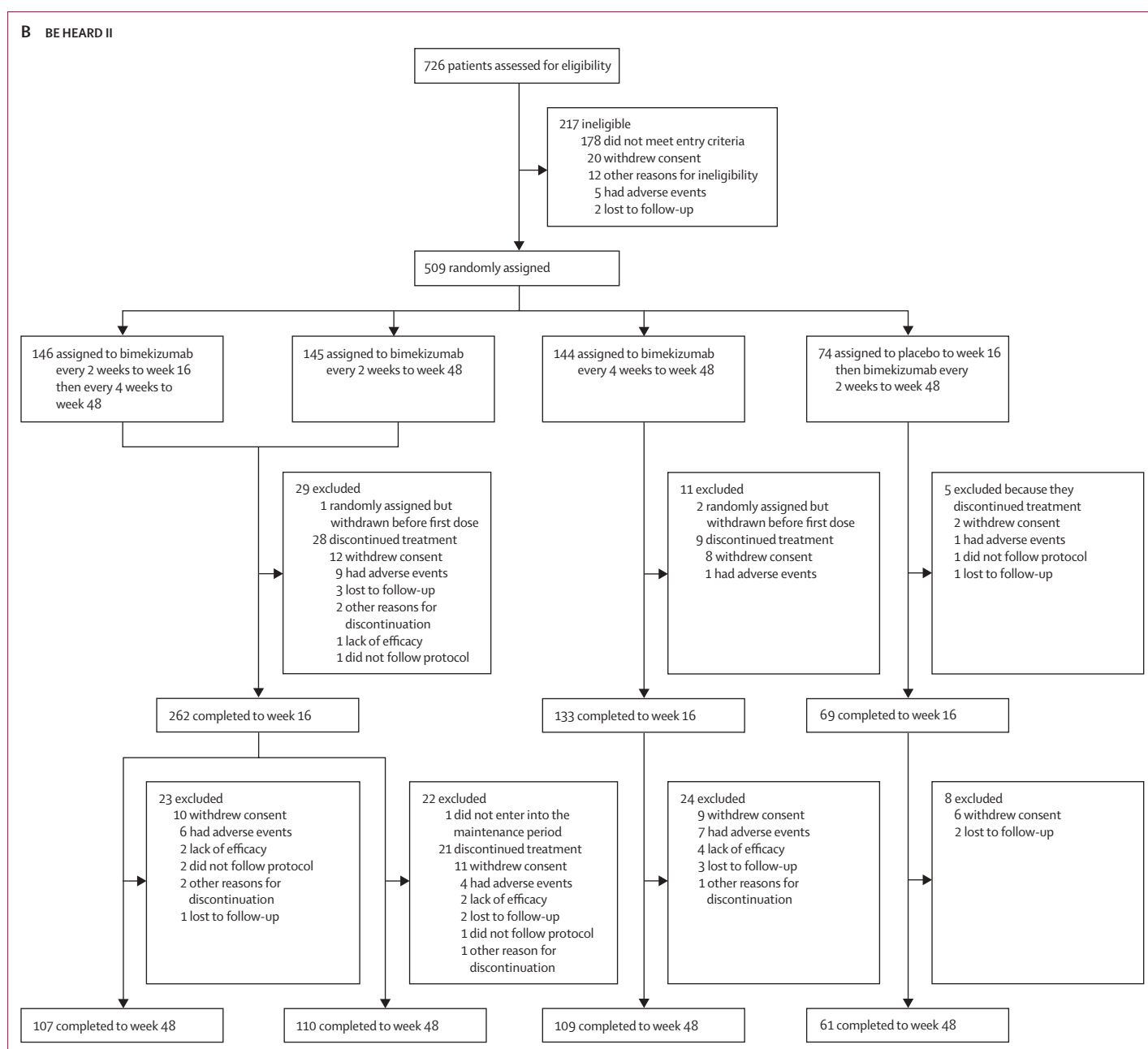


Figure 1: Trial profiles

The BE HEARD I (A) and BE HEARD II (B) trial profiles to week 48. Patients randomly assigned to bimekizumab 320 mg every 2 weeks at baseline were pooled to week 16, per prespecified analysis. For the maintenance treatment period, the four dose regimens are presented separately. *In BE HEARD I, one participant did not pass screening and did not complete the study termination CRF page, as such no primary reason for screen failure is listed for this participant and they are included in the category of other reasons for ineligibility.

observed in bimekizumab-treated groups than in the placebo group. In BE HEARD I, 138 (48%) of 289 patients in the group who received bimekizumab every 2 weeks versus 21 (29%) of 72 patients in the placebo group met HiSCR50 (OR 2.23 [97.5% CI 1.16–4.31]; $p=0.0060$; table 2). In BE HEARD II, 151 (52%) of 291 patients in the group who received bimekizumab every 2 weeks versus 24 (32%) of 74 patients in the placebo group met HiSCR50

(2.29 [1.22–4.29]; $p=0.0032$). HiSCR50 in patients who received bimekizumab every 2 weeks to week 16 was rapidly reached across both trials (figure 2A, B). Responses were observed as early as week 4 and were maintained or increased to week 48 (figure 3A, B).

The primary outcome was also met in the group who received bimekizumab every 4 weeks in BE HEARD II. 77 (54%) of 144 patients versus 24 (32%) of 74 patients in

	BE HEARD I			BE HEARD II		
	Bimekizumab 320 mg every 2 weeks* (n=289)	Bimekizumab 320 mg every 4 weeks (n=144)	Placebo (n=72)	Bimekizumab 320 mg every 2 weeks* (n=291)	Bimekizumab 320 mg every 4 weeks (n=144)	Placebo (n=74)
Demographics						
Age, years	36.0 (26.0–46.0)	35.0 (27.0–45.0)	33.5 (26.0–46.0)	35.0 (27.0–45.0)	33.0 (26.0–42.5)	37.0 (28.0–47.0)
Age group, years						
<40 years	174 (60%)	93 (65%)	45 (63%)	180 (62%)	97 (67%)	46 (62%)
40 years to <65 years	109 (38%)	50 (35%)	26 (36%)	107 (37%)	45 (31%)	24 (32%)
≥65 years	6 (2%)	1 (<1%)	1 (1%)	4 (1%)	2 (1%)	4 (5%)
Sex						
Female	176 (61%)	98 (68%)	44 (61%)	150 (52%)	77 (54%)	31 (42%)
Male	113 (39%)	46 (32%)	28 (39%)	141 (48%)	67 (46%)	43 (58%)
Body weight, kg	97.2 (25.4)	102.7 (24.7)	94.6 (24.8)	95.4 (24.2)	95.3 (22.0)	100.3 (23.7)
BMI, kg/m ²	33.4 (8.3)	35.4 (8.1)	32.4 (7.8)	32.0 (8.0)	32.2 (7.5)	33.8 (8.7)
Smoking status						
Current	127 (44%)	53 (37%)	37 (51%)	134 (46%)	73 (51%)	38 (51%)
Former†	43 (15%)	28 (19%)	7 (10%)	49 (17%)	14 (10%)	10 (14%)
Race						
White	233 (81%)	105 (73%)	55 (76%)	232 (80%)	119 (83%)	64 (86%)
Black	41 (14%)	21 (15%)	8 (11%)	22 (8%)	13 (9%)	5 (7%)
Asian	2 (<1%)	3 (2%)	3 (4%)	22 (8%)	7 (5%)	5 (7%)
Previous use of biological therapy‡	76 (26%)	31 (22%)	19 (26%)	41 (14%)	16 (11%)	10 (14%)
Baseline disease characteristics						
Disease duration, years	5.7 (3.1–12.0)	5.6 (2.6–11.9)	8.7 (4.5–15.4)	4.9 (2.1–10.4)	4.2 (1.9–7.8)	4.8 (1.3–12.3)
Abscess and inflammatory nodule count	15.3 (13.5)	17.8 (25.3)	15.0 (11.9)	16.7 (15.5)	17.6 (15.4)	13.9 (7.8)
Abscess count	3.7 (6.1)	4.5 (8.4)	2.9 (6.6)	3.3 (5.9)	3.5 (5.0)	2.4 (2.8)
Inflammatory nodule count	11.6 (11.4)	13.3 (22.4)	12.2 (10.0)	13.4 (12.2)	14.1 (13.3)	11.4 (6.7)
Draining tunnel count	4.0 (4.9)	3.8 (4.9)	3.2 (4.0)	3.6 (4.0)	2.8 (3.1)	3.5 (3.7)
Hurley stage						
II	149 (52%)	71 (49%)	34 (47%)	177 (61%)	89 (62%)	45 (61%)
III	140 (48%)	73 (51%)	38 (53%)	114 (39%)	55 (38%)	29 (39%)
DLQI total score	11.5 (6.6)	12.8 (7.6)	12.4 (8.0)	10.6 (6.5)	10.5 (7.0)	11.9 (6.1)
Concomitant antibiotic use	27 (9%)	8 (6%)	5 (7%)	30 (10%)	10 (7%)	6 (8%)
HSSDD worst skin pain score	5.5 (2.5)	5.9 (2.6)	6.0 (2.5)	5.3 (2.4)	5.3 (2.5)	5.0 (2.4)

Data are median (IQR), mean (SD), or n (%). Percentages might not add up to 100% due to rounding. DLQI=Dermatology Life Quality Index. HSSDD=hidradenitis suppurativa symptom daily diary. *Data were pooled for all patients randomly assigned to bimekizumab 320 mg every 2 weeks for the first 16 weeks. †Patients were included in the former smoker category if they had been a smoker at any previous point. ‡A full list of excluded prior biological agents is shown in appendix 2 (pp 43–45, 167–169).

Table 1: Baseline demographics and disease characteristics in BE HEARD I and II

the placebo group met HiSCR50 (OR 2.42 [97.5% CI 1.22–4.80]; $p=0.0038$; table 2). In BE HEARD I, the primary outcome was not met for patients treated with bimekizumab every 4 weeks. 65 (45%) of 144 patients versus 21 (29%) of 72 patients in the placebo group (2.00 [0.98–4.09]; $p=0.030$) met HiSCR50. Similarly to the group that received bimekizumab every 2 weeks, the response in the group that received bimekizumab every 4 weeks was rapidly reached in both trials (figure 2A, B), with initial improvements observed by week 4 and maintained or increased over 48 weeks of treatment (figure 3A, B).

For both trials, the key secondary outcome of HiSCR75 at week 16 was met in the group that received bimekizumab every 2 weeks. In BE HEARD I, 97 (33%) of 289 patients met HiSCR75 versus 13 (18%) of 72 patients treated with placebo (table 2). In BE HEARD II, 104 (36%) of 291 patients who received bimekizumab every 2 weeks versus 12 (16%) of 74 patients treated with placebo met HiSCR75. The group that received bimekizumab every 2 weeks met HiSCR75 rapidly (by week 4) across both BE HEARD I and II (figure 2C, D), with the proportions of patients who met HiSCR75 maintained or increased to week 48 (figure 3C, D).

	BE HEARD I			BE HEARD II		
	Bimekizumab 320 mg every 2 weeks* (n=289)	Bimekizumab 320 mg every 4 weeks (n=144)	Placebo (n=72)	Bimekizumab 320 mg every 2 weeks* (n=291)	Bimekizumab 320 mg every 4 weeks (n=144)	Placebo (n=74)
Primary efficacy endpoint						
HiSCR50†‡	138 (48%)	65 (45%)	21 (29%)	151 (52%)	77 (54%)	24 (32%)
Bimekizumab every 2 weeks vs placebo	2.23 (1.16 to 4.31); p=0.0060§	2.29 (1.22 to 4.29); p=0.0032§
Bimekizumab every 4 weeks vs placebo	..	2.00 (0.98 to 4.09); p=0.030¶	2.42 (1.22 to 4.80); p=0.0038§	..
Ranked secondary endpoints						
HiSCR75†‡	97 (33%)	36 (25%)	13 (18)	104 (36%)	49 (34%)	12 (16%)
Bimekizumab every 2 weeks vs placebo	2.18 (1.02 to 4.64); p=0.021§	3.01 (1.37 to 6.58); p=0.0016§
Bimekizumab every 4 weeks vs placebo	..	1.42 (0.62 to 3.26); p=0.35¶	2.72 (1.18 to 6.27); p=0.0071§	..
Flare†	NA	NA	NA	84 (29%)	34 (24%)	21 (28%)
Bimekizumab every 2 weeks vs placebo	NA	NA	NA	1.05 (0.54 to 2.04); p=0.87¶
Bimekizumab every 4 weeks vs placebo	NA	NA	NA	..	0.80 (0.38 to 1.68); p=0.50¶	..
DLQI total score change from baseline**††	-5.0 (0.4)	-5.5 (0.5)	-2.7 (0.7)	-4.5 (0.3)	-4.1 (0.4)	-3.1 (0.6)
Bimekizumab every 2 weeks vs placebo, least-squares-mean difference (97.5% CI); p value	-2.68 (-4.39 to -0.97); p=0.0005§	-2.31 (-3.71 to -0.91); p=0.0002¶
Bimekizumab every 4 weeks vs placebo, least-squares-mean difference (97.5% CI); p value	..	-2.57 (-4.47 to -0.68); p=0.0024¶	-2.39 (-3.92 to -0.87); p=0.0004¶	..
HSSDD worst skin pain score change from baseline**‡‡	-1.9 (0.2)	-1.7 (0.2)	-1.1 (0.2)	-1.9 (0.1)	-1.7 (0.2)	-0.4 (0.3)
Bimekizumab every 2 weeks vs placebo, least-squares-mean difference (97.5% CI); p value	-1.19 (-2.05 to -0.32); p=0.0022§	-1.27 (-1.98 to -0.55); p<0.0001¶
Bimekizumab every 4 weeks vs placebo, least-squares-mean difference (97.5% CI); p value	..	-0.55 (-1.52 to 0.42); p=0.20¶	-0.90 (-1.68 to -0.11); p=0.010¶	..
HSSDD worst skin pain response†§§	61 (32%)	23 (22%)	7 (15%)	66 (32%)	31 (29%)	5 (11%)
Bimekizumab every 2 weeks vs placebo	2.76 (0.91 to 8.36); p=0.041¶	3.76 (1.19 to 11.87); p=0.010¶
Bimekizumab every 4 weeks vs placebo	..	1.62 (0.49 to 5.35); p=0.37¶	3.27 (0.97 to 11.00); p=0.028¶	..

Data are OR (97.5% CI); p value or mean (SE), unless otherwise specified for the randomised set. ORs are presented for binary variables and least-squares-mean difference presented for continuous variables. For multiply imputed binary variables, the rounded average number of patients with response based on 100 imputations is reported. DLQI=Dermatology Life Quality Index. HiSCR=hidradenitis suppurativa clinical response. HiSCR50=reduction in total abscess and inflammatory nodule count of at least 50% from baseline with no increase from baseline in abscess or draining tunnel count. HiSCR75=reduction in total abscess and inflammatory nodule count of at least 75% from baseline with no increase from baseline in abscess or draining tunnel count. HSSDD=hidradenitis suppurativa symptom daily diary. NA=not applicable. OR=odds ratio. *Data were pooled for all patients randomly assigned to bimekizumab 320 mg every 2 weeks for the first 16 weeks. †Data were imputed by means of a modified non-responder imputation (all-antibiotics): patients who took any systemic antibiotic (new or increased dose) or who discontinued due to an adverse event or absence of efficacy were treated as non-responders (or treated as experiencing flare for the flare endpoint) at all subsequent visits. Other missing data were imputed via multiple imputation (primary, pre-specified analysis method). ‡p values (from Wald tests) for adjusted responder rates obtained from logistic regression with treatment, Hurley stage at baseline, and baseline antibiotic use (and analgesic use for pain response only) as factors. §Statistically significant per the statistical hierarchy. ¶p value calculated based on statistical testing methodology, had the given bimekizumab regimen succeeded at hierarchical testing. ||Flare by week 16 was defined as at least one occurrence of flare between baseline and up to week 16, in which flare was defined as at least a 25% increase in abscess and inflammatory nodule count with an increase of at least two abscess and inflammatory nodules relative to baseline. Flare was not a secondary endpoint in BE HEARD I, so these cells have been marked with NA. **Data were imputed using multiple imputation (all-antibiotics): patients who discontinued study treatment due to absence of efficacy or adverse events, or who received any systemic antibiotics during the study (new or increased dose), were set to missing and subsequently imputed using multiple imputation. All other missing data were also imputed using multiple imputation. ††p values based on an ANCOVA with fixed effects of treatment, Hurley stage at baseline, baseline antibiotic use, and baseline DLQI total score as covariates. ‡‡p values based on an ANCOVA with fixed effects of treatment, Hurley stage at baseline, baseline antibiotic use, analgesic use, and baseline HSSDD worst skin pain score as covariates. §§Pain response was defined as an improvement from baseline in HSSDD weekly worst skin pain score of at least 3 points among patients with a baseline score of 3 or higher.

Table 2: Primary and key secondary efficacy outcomes in BE HEARD I and BE HEARD II at week 16

In BE HEARD II, HiSCR75 at week 16 was met in the group who received bimekizumab every 4 weeks in 49 (34%) of 144 patients versus 12 (16%) of 74 patients treated with placebo (table 2). In the group who received bimekizumab every 4 weeks in BE HEARD I, 36 (25%) of 144 met HiSCR75 versus 13 (18%) of 72 patients treated with placebo (not statistically significant).

Patients switching from placebo to bimekizumab every 2 weeks showed rapid improvements in HiSCR50 and HiSCR75 in both trials after switching to bimekizumab, with responses increased to week 48 (figure 3).

On the basis of mNRI (hidradenitis suppurativa-antibiotics), HiSCR50 at week 48 in BE HEARD I was met by 87 (61%) of 143 patients receiving bimekizumab

every 2 weeks, 90 (61%) of 146 patients receiving bimekizumab every 2 weeks followed by every 4 weeks, 76 (53%) of 144 patients receiving bimekizumab every 4 weeks, and 33 (45%) of 72 patients receiving placebo followed by bimekizumab every 2 weeks (appendix 1 p 29). In BE HEARD II, HiSCR50 at week 48 was met by 88 (61%) of 145 patients receiving bimekizumab every 2 weeks, 93 (64%) of 146 patients receiving bimekizumab every 2 weeks followed by every 4 weeks, 91 (63%) of 144 patients receiving bimekizumab every 4 weeks, and 50 (68%) of 74 patients receiving placebo followed by bimekizumab every 2 weeks. HiSCR75 results calculated by means of mNRI (hidradenitis suppurativa-antibiotics) are also reported in appendix 1 (p 29). HiSCR90 and HiSCR100 responses by week 16 improved to week 48 with bimekizumab treatment across BE HEARD I and BE HEARD II (appendix 1 pp 19, 30–31).

In BE HEARD II, no significant differences in flare were detected between the treatment and placebo groups (table 2). Across both studies, reductions from baseline in abscess and inflammatory nodules and draining tunnel counts were observed by week 4 across bimekizumab treatment regimens, with counts numerically lower among patients treated with bimekizumab versus placebo (appendix 1 pp 20, 32–33).

Patients treated with bimekizumab had greater improvements in patient-reported outcomes measuring effect of hidradenitis suppurativa on health-related quality of life and core hidradenitis suppurativa symptoms, including worst skin pain, compared with the placebo group at week 16. Across BE HEARD I and BE HEARD II, both bimekizumab treatment regimens had numerically greater improvements (ie, decreased score) in DLQI and HSSDD worst skin pain (ranked key secondary outcomes) versus placebo at week 16, with clinically meaningful improvements observed (table 2). Patients treated with bimekizumab had rapid improvements in HSSDD worst skin pain, as early as week 2, that were maintained over the 16 weeks in which HSSDD was assessed (appendix 1 p 34).

Bimekizumab was well tolerated at both dosing regimens; the safety profile was consistent with the phase 2 study of bimekizumab in hidradenitis suppurativa and phase 3 studies for other diseases.^{17,18,22–24} During the initial placebo-controlled, 16-week treatment period, the frequency of treatment-emergent adverse events was generally similar across groups given bimekizumab every 2 weeks, those given bimekizumab every 4 weeks, and those who received placebo in both trials (table 3; appendix 1 pp 21–22). Serious or severe treatment-emergent adverse events were infrequent in both groups but numerically superior in patients treated with bimekizumab than in patients treated with placebo. No clinically meaningful differences in the incidence of serious or severe treatment-emergent adverse events or of treatment-emergent adverse events leading to discontinuation were detected between bimekizumab treatment

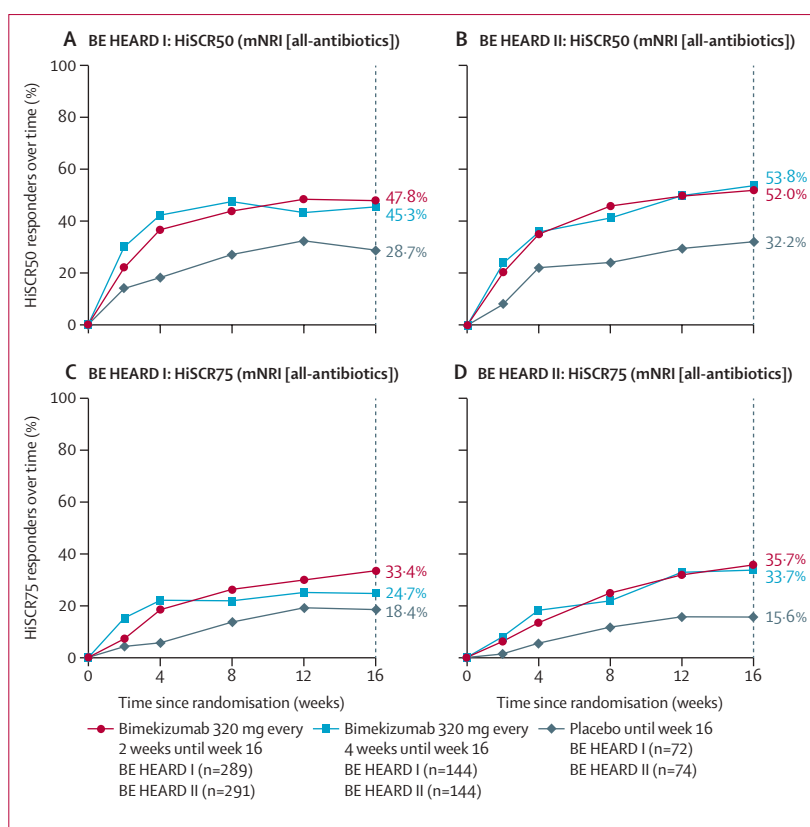


Figure 2: HiSCR responses from randomisation to week 16

The rates of HiSCR50 (A, B) and HiSCR75 (C, D) from randomisation to week 16 in BE HEARD I and BE HEARD II. mNRI (all-antibiotics): patients who took any systemic antibiotic (new or increased dose) or who discontinued due to an adverse event or absence of efficacy were treated as non-responders at all subsequent visits. Other missing data were imputed via multiple imputation (primary, pre-specified analysis method). HiSCR=hidradenitis suppurativa clinical response. HiSCR50=reduction in total abscess and inflammatory nodule count of at least 50% from baseline with no increase from baseline in abscess or draining tunnel count. HiSCR75=reduction in total abscess and inflammatory nodule count of at least 75% from baseline with no increase from baseline in abscess or draining tunnel count. mNRI=modified non-responder imputation.

groups. Similarly, during weeks 0–48, no clinically meaningful differences were noted in the incidence of discontinuations due to treatment-emergent adverse events between treatment groups, and serious treatment-emergent adverse events were infrequent and occurred at similar rates between bimekizumab treatment groups in both trials (appendix 1 pp 23–24). The most frequently reported treatment-emergent adverse events to week 48 were hidradenitis in both BE HEARD I and BE HEARD II (related to hidradenitis suppurativa worsening), in addition to SARS-CoV-2 infection and diarrhoea in BE HEARD I, and oral candidiasis and headache in BE HEARD II.

Across 48 weeks, at least one treatment-emergent adverse event occurred in 425 (86%) patients who received bimekizumab in BE HEARD I and 412 (82%) in BE HEARD II, with similar rates across dosing groups (table 3; appendix 1 pp 23–24). One death, caused by congestive heart failure in a patient with cardiovascular history treated with bimekizumab every 2 weeks in BE HEARD I, was reported across the two

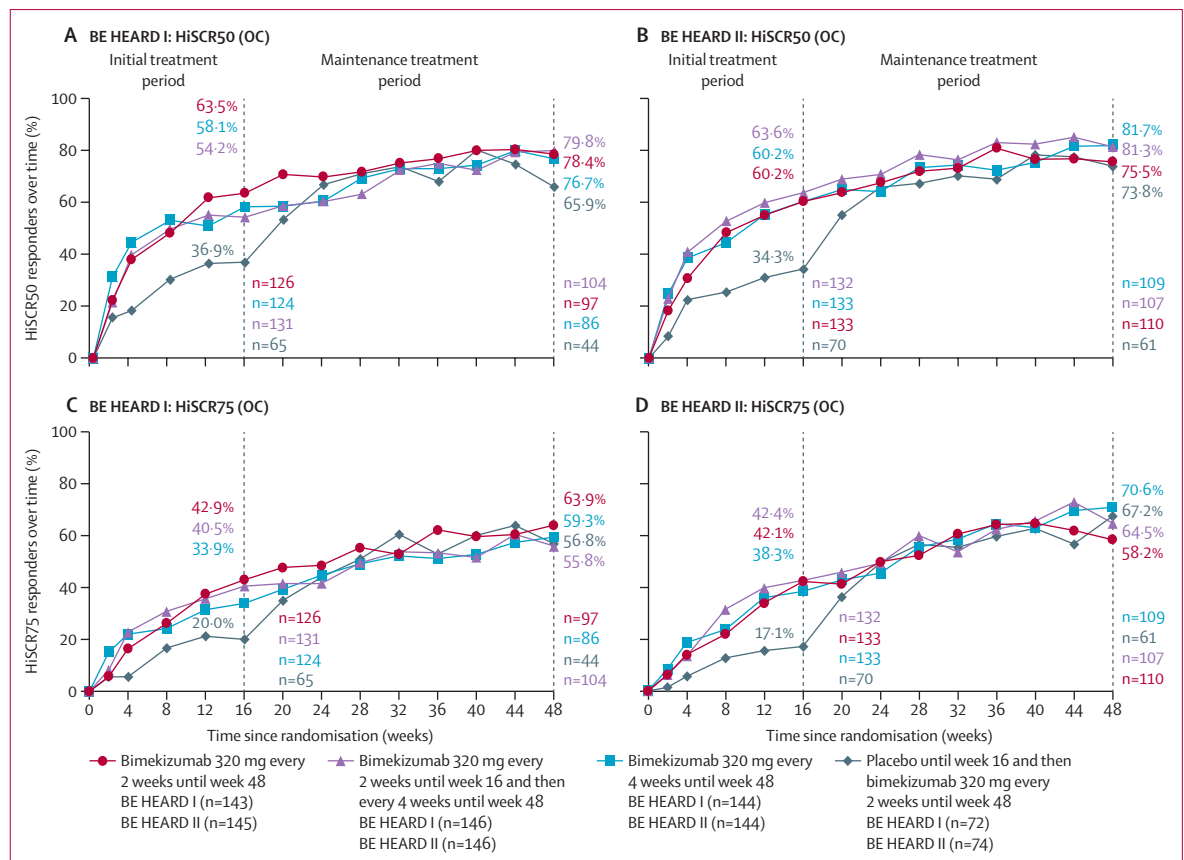


Figure 3: HiSCR responses from randomisation to week 48

The rates of HiSCR50 (A, B) and HiSCR75 (C, D) from randomisation to week 48 in BE HEARD I and BE HEARD II. Data are shown for observed cases (ie, all available data after an intercurrent event were summarised as recorded in the database, and all missing data were left missing). HiSCR=hidradenitis suppurativa clinical response. HiSCR50=reduction in total abscess and inflammatory nodule count of at least 50% from baseline with no increase from baseline in abscess or draining tunnel count. HiSCR75=reduction in total abscess and inflammatory nodule count of at least 75% from baseline with no increase from baseline in abscess or draining tunnel count.

trials and considered unrelated to bimekizumab treatment by the investigator. Study discontinuation due to treatment-emergent adverse events was similar across both trials and treatment groups, occurring in 40 (8%) patients treated with bimekizumab in BE HEARD I and 27 (5%) patients treated with bimekizumab in BE HEARD II, to 48 weeks (table 3; appendix 1 pp 23–24).

By week 48, fungal infection events occurred in 112 (23%) patients treated with bimekizumab in BE HEARD I and 124 (25%) of those treated with bimekizumab in BE HEARD II, with incidence generally similar across bimekizumab treatment groups (table 3; appendix 1 pp 23–24), and lower with placebo to week 16 (table 3). Of these fungal events, *Candida* spp infections were most common. Oral candidiasis occurred in 47 (10%) patients treated with bimekizumab BE HEARD I and 64 (13%) of those treated with bimekizumab in BE HEARD II (table 3). Most cases of oral candidiasis were mild to moderate, resolved following standard anti-fungal therapy, and did not lead to discontinuation.

Hypersensitivity reactions, mostly related to the skin (eg, dermatitis and eczema), occurred in 105 (21%) patients treated with bimekizumab in BE HEARD I and 84 (17%) of those treated with bimekizumab in BE HEARD II (table 3). No anaphylaxis events related to bimekizumab occurred. Most hypersensitivity reactions were mild to moderate and did not lead to discontinuation; one serious case occurred which led to discontinuation.

Three patients treated with bimekizumab in BE HEARD I and four in BE HEARD II had adjudicated definite or probable inflammatory bowel disease, all of which were of new onset. Of these, one (BE HEARD I) and three (BE HEARD II) patients with the condition discontinued treatment. In the six patients in BE HEARD I and two in BE HEARD II with a previous history of inflammatory bowel disease, no flares were reported. Incidences of neutropenia, adjudicated major adverse cardiovascular events, and malignancies were low to week 48 (table 3). Of adjudicated hepatic events, no elevations of aspartate aminotransferase or alanine

	Initial treatment period only (weeks 0–16)						Initial and maintenance treatment period (weeks 0–48)	
	BE HEARD I			BE HEARD II			BE HEARD I	BE HEARD II
	Bimekizumab 320 mg every 2 weeks* (n=286); 100 PY=0.87	Bimekizumab 320 mg every 4 weeks (n=143); 100 PY=0.43	Placebo (n=72); 100 PY=0.22	Bimekizumab 320 mg every 2 weeks* (n=290); 100 PY=0.88	Bimekizumab 320 mg every 4 weeks (n=142); 100 PY=0.44	Placebo (n=74); 100 PY=0.23	Bimekizumab total (n=494); 100 PY=3.99	Bimekizumab total (n=501); 100 PY=4.14
Any TEAE	192 (67%)	94 (66%)	48 (67%)	187 (64%)	73 (51%)	42 (57%)	425 (86%)	412 (82%)
Serious TEAE	6 (2%)	4 (3%)	0	9 (3%)	3 (2%)	0	40 (8%)	24 (5%)
Discontinuation due to TEAE	10 (3%)	6 (4%)	1 (1%)	12 (4%)	3 (2%)	0	40 (8%)	27 (5%)
Drug-related TEAE	84 (29%)	37 (26%)	12 (17%)	105 (36%)	38 (27%)	8 (11%)	227 (46%)	217 (43%)
Severe TEAE	8 (3%)	3 (2%)	0	12 (4%)	5 (4%)	2 (3%)	42 (9%)	39 (8%)
Deaths	0	0	0	0	0	0	1 (<1%)	0
Most common TEAEs								
Hidradenitis	19 (7%)	12 (8%)	10 (14%)	25 (9%)	13 (9%)	5 (7%)	96 (19%)	90 (18%)
Coronavirus infection	9 (3%)	2 (1%)	2 (3%)	11 (4%)	3 (2%)	0	71 (14%)	36 (7%)
Oral candidiasis	17 (6%)	2 (1%)	0	24 (8%)	5 (4%)	0	47 (10%)	64 (13%)
Diarrhoea	18 (6%)	12 (8%)	1 (1%)	18 (6%)	5 (4%)	6 (8%)	49 (10%)	36 (7%)
Headache	22 (8%)	8 (6%)	3 (4%)	18 (6%)	7 (5%)	7 (9%)	43 (9%)	43 (9%)
TEAEs of interest								
Infections and infestations	98 (34%)	52 (36%)	18 (25%)	95 (33%)	39 (27%)	12 (16%)	301 (61%)	277 (55%)
Serious infections	1 (<1%)	0	0	0	0	0	11 (2%)	5 (1%)
Opportunistic infections†	1 (<1%)	1 (<1%)	0	0	1 (<1%)	0	8 (2%)	4 (<1%)
Fungal infections	34 (12%)	17 (12%)	1 (1%)	41 (14%)	18 (13%)	0	112 (23%)	124 (25%)
Candida spp infections	22 (8%)	7 (5%)	0	26 (9%)	15 (11%)	0	67 (14%)	86 (17%)
Oral candidiasis	17 (6%)	2 (1%)	0	24 (8%)	5 (4%)	0	47 (10%)	64 (13%)
Neutropenia	0	0	0	0	0	0	1 (<1%)	0
Hypersensitivity reaction (SMQ, narrow)‡	30 (10%)	12 (8%)	4 (6%)	32 (11%)	9 (6%)	1 (1%)	105 (21%)	84 (17%)
Dermatitis and eczema	14 (5%)	6 (4%)	3 (4%)	21 (7%)	8 (6%)	1 (1%)	62 (13%)	60 (12%)
Serious hypersensitivity reaction	0	0	0	0	0	0	0	1 (<1%)
Adjudicated suicidal ideation and behaviour	0	2 (1%)	0	1 (<1%)	0	0	5 (1%)	1 (<1%)
Adjudicated major adverse cardiovascular events	0	0	0	0	0	0	3 (<1%)	0
Hepatic events§	8 (3%)	2 (1%)	4 (6%)	6 (2%)	3 (2%)	0	25 (5%)	19 (4%)
>5 times the ULN elevation of AST/ALT¶	3/284 (1%)	0/140	0/71	0/288	0	0/73	4/489 (<1%)	4/499 (<1%)
Malignancies	0	0	0	1 (<1%)	0	0	1 (<1%)	3 (<1%)
Definite or probable adjudicated inflammatory bowel disease	0	1 (<1%)	0	1 (<1%)	2 (1%)	0	3 (<1%)	4 (<1%)

Data are n (%) and all denominators are indicated in the column headers, except where otherwise specified. Adverse events in the safety set (weeks 0–16) and active medication set (weeks 0–48), as per MedDRA (version 19.0). ALT=alanine aminotransferase. AST=aspartate aminotransferase. MedDRA=Medical Dictionary for Regulatory Activities. PY=patient-year. SMQ=Standardised MedDRA Query. TEAE=treatment-emergent adverse event. ULN=upper limit of normal. *Data were pooled for all patients randomly assigned to bimekizumab 320 mg every 2 weeks for the first 16 weeks. †Opportunistic infections were localised mucocutaneous events, as defined by internal company conventions. ‡Using the narrow SMQ definition of hypersensitivity reaction events. §The hepatic events category includes events in the SMQ drug-related hepatic disorders comprehensive search SMQ, excluding the two sub-SMQs of benign liver neoplasms (including cysts and polyps) SMQ and malignant and unspecified liver neoplasms SMQ. ¶No elevations of greater than five times the ULN were adjudicated to be highly likely or definitely related to bimekizumab.

Table 3: Adverse events per 100 patient-years to week 16 and week 48 in BE HEARD I and BE HEARD II

aminotransferase greater than five times the upper limit of normal were adjudicated to be highly likely or definitely related to bimekizumab. Both studies exhaustively monitored and collected patient data related to suicidality and depression through questionnaires. Overall, five patients in BE HEARD I and one in BE HEARD II had adjudicated

suicidal ideation and behaviour, with no events of completed suicide (table 3).

Discussion

In the BE HEARD I and BE HEARD II trials, bimekizumab significantly reduced signs of disease in

patients with moderate-to-severe hidradenitis suppurativa compared with placebo. The primary outcome of **HiSCR50** was significantly higher in the group given **bimekizumab every 2 weeks** than in the placebo group at week 16 in both trials, with responses maintained (and improved for HiSCR75 and HiSCR90) to week 48. Efficacy of bimekizumab in patients treated with bimekizumab every 4 weeks was significantly higher than placebo in BE HEARD II trial only, although similar response rates were observed with every 2 weeks dosing in BE HEARD I.

HiSCR50 was created for the seminal adalimumab trials and helped establish the concept of biological therapy in hidradenitis suppurativa.²⁰ It has since been used as the primary outcome in numerous hidradenitis suppurativa clinical studies; however, because this outcome represents an improvement in abscess and inflammatory nodule count of only 50%, as research advances, studies have begun measuring higher thresholds of clinical improvement reflecting deeper clinical responses in patients with hidradenitis suppurativa.²⁵ In these two trials, bimekizumab treatment produced rapid and deep clinical responses, with greater proportions of patients treated with bimekizumab meeting the more stringent outcomes of HiSCR75 and HiSCR90 versus placebo as early as week 4, and responses maintained or improved over 48 weeks of treatment.

Other phase 3 hidradenitis suppurativa programmes suggest that continued treatment beyond a 12–16 week primary outcome measure could lead to further improvement.⁹ With bimekizumab treatment, clinical responses were sustained or improved over 48 weeks, with rapid improvements (within 4 weeks) observed in patients who switched from placebo to bimekizumab from week 16 onwards. Given the limited efficacy of current therapeutic options available for patients with hidradenitis suppurativa, the rapid, deep, and maintained clinical response improvements offered by bimekizumab provide a potential additional treatment option; however, because of the heterogeneous nature of the disease, it must be acknowledged that clinical response might be affected by an individual patient's underlying disease aetiology, and thus IL-17F and IL-17A should not be considered as the only drivers of disease among patients with hidradenitis suppurativa.^{5,26}

Across both trials, bimekizumab showed improvements in patient-reported outcomes alongside clinical improvements. Clinically meaningful improvements in DLQI were observed at week 16 in bimekizumab treatment groups, but not with placebo. Other useful tools include patient-reported items such as the pain index and HSSDD.^{19,27} Pain was the highest ranked item in the hidradenitis suppurativa Core Outcomes Set International Collaboration and is an important, albeit challenging and evolving outcome measure in hidradenitis suppurativa research.⁶ Rapid and clinically meaningful improvements in skin pain were observed in

BE HEARD I and II, with improvements in HSSDD worst skin pain versus placebo at week 16 observed in patients who received bimekizumab every 2 weeks. Further data on signs of disease using the International hidradenitis suppurativa Severity Score System (IHSS4) and hidradenitis suppurativa-specific health-related quality of life using the hidradenitis suppurativa Quality of Life (HiSQoL) were collected in the BE HEARD programme and will be published in a dedicated manuscript.^{28,29}

Systemic antibiotics are frequently used as first-line therapy for flares or as cotreatment for hidradenitis suppurativa.³⁰ They are commonly evaluated as adjuvant therapy in clinical trials of hidradenitis suppurativa, although methodology to calculate efficacy in the presence of systemic antibiotic use is not standardised. In BE HEARD I and II, systemic antibiotic use consistent with hidradenitis suppurativa treatment guidelines was permitted as rescue therapy for flaring disease.³⁰ Patients who took any systemic antibiotic for any indication, including for non-hidradenitis suppurativa indications, were treated as non-responders at subsequent visits for the primary analysis at week 16 (mNRI [all-antibiotics]), which is likely to underestimate the efficacy of bimekizumab for hidradenitis suppurativa, given that patients responding well to bimekizumab but who receive systemic antibiotic therapy for reasons not related to hidradenitis suppurativa are considered non-responders. In the supportive post-hoc analyses, patients treated with systemic antibiotics identified as rescue therapy for hidradenitis suppurativa by the investigator were treated as non-responders at subsequent visits for the evaluation of outcomes over time (mNRI [hidradenitis suppurativa-antibiotics]). Despite the stringent nature of the analysis, the studies showed high levels of maintained response over time. By contrast, long-term (52-week) data reported from the SUNSHINE and SUNRISE trials of secukinumab in hidradenitis suppurativa reported observed case data only from weeks 18–52.⁹

The safety profile of bimekizumab in BE HEARD I and II was consistent with other indications, and with the selective IL-17A inhibitor, secukinumab, which is approved for use in hidradenitis suppurativa.^{9,17,22–24} Moreover, over 48 weeks, the most common treatment-emergent adverse events were similar between bimekizumab treatment groups. Candidiasis and inflammatory bowel disease are associated with IL-17 inhibition.^{31,32} Candidiasis incidence was higher among patients who received bimekizumab treatment than in those who received placebo. Oral candidiasis events were generally mild to moderate, resolved following standard therapy, and did not lead to discontinuation. Across the 1014 patients randomly assigned in BE HEARD I and II, incidence of safety topics of interest, such as inflammatory bowel disease and suicidal ideation and behaviour, were aligned with study population expectations.⁵ Adjudicated definite or probable inflammatory bowel disease occurred

in three patients on bimekizumab in BE HEARD I and four in BE HEARD II, with no flares in patients with a previous history of inflammatory bowel disease. Although paradoxical hidradenitis suppurativa reactions have been reported following treatment with both anti-TNF and anti-IL-17A-only agents,³³ as per other hidradenitis suppurativa clinical trials, hidradenitis is recorded as an adverse event in the BE HEARD I and II trials because of the flaring nature of the disease.

Limitations of the BE HEARD I and II trials include the relatively short initial 16-week placebo-controlled period, which might affect the interpretability of later efficacy results, and the absence of an active comparator across 48 weeks of treatment. In addition, evaluation of treatment efficacy in the presence of rescue systemic antibiotic use poses a challenge; the methods used to calculate efficacy rates under these conditions has not yet been standardised across trials in hidradenitis suppurativa. In clinical trials of patients with hidradenitis suppurativa, a patient's underlying disease severity or multifactorial aetiology of underlying disease could lead to variability in observed efficacy and subsequent interpretation of results.²⁶ Subsequent studies are warranted, including collection of real-world evidence for bimekizumab use in patients with hidradenitis suppurativa, considerations for precision medicine approaches in hidradenitis suppurativa, network meta-analyses to inform clinical decision making, and surveillance of concomitant medical and surgical therapy use. Hidradenitis suppurativa head-to-head comparator studies are scarce. Although numerically greater proportions of patients met HiSCR50 by week 48 in BE HEARD I and II than in other long-term phase 3 trials of hidradenitis suppurativa (using observed case analysis), further research is needed to formally compare these outcomes between biologic therapies because of the inevitable heterogeneity in trial populations and differing analysis methods used across studies.^{9,10}

These two phase 3 trials demonstrated that bimekizumab was well tolerated by patients with moderate-to-severe hidradenitis suppurativa and that it produced rapid, deep, and maintained clinically meaningful improvements in physician-assessed and patient-reported outcome measures to week 48. These data support the use of bimekizumab as a promising new therapeutic option for patients with moderate-to-severe hidradenitis suppurativa.

Contributors

All authors contributed to study conception and design, contributed to data analysis and interpretation, and contributed to drafting and revision of the manuscript for important intellectual content. All authors had access to and verified all the included data in the study and had final responsibility for the decision to submit for publication.

Declaration of interests

ABK reports grants paid to the institution from AbbVie, Amgen, Anaptys Bio, Arista, Bristol Myers Squibb, Eli Lilly and Company, Incyte, Janssen, MoonLake Immunotherapeutics, Novartis, Pfizer, Prometheus, Sonoma Bio, and UCB Pharma; consulting fees from AbbVie, Alumis, Bayer, Boehringer Ingelheim, Eli Lilly and Company, Janssen, MoonLake Immunotherapeutics, Novartis, Pfizer, Prionviant, Sonoma Bio, Sanofi,

Target RWE, UCB Pharma, Union, and Ventyx; and serves on the board of directors of Almirall. GBEJ reports honoraria for participation on advisory boards from AbbVie, Boehringer Ingelheim, Chemocentryx, Incyte, Janssen-Cilag, LEO Pharma, Novartis, and UCB Pharma; and acted as an investigator for AbbVie, CJSL, InflaRx, Janssen-Cilag, LEO Pharma, Novartis, Regeneron, Sanofi, and UCB Pharma. CJS acted as an investigator for AbbVie, Chemocentryx, GSK, Incyte, InflaRx, Novartis, and UCB Pharma; reports consultancy fees from AbbVie, Alumis, InflaRx, Incyte, Logical Images, Sonoma Biotherapeutics, and UCB Pharma; and acted as a speaker for AbbVie and Novartis. JSK reports personal fees from AbbVie, ChemoCentryx, CJSL Behring, DermTech, Incyte, Insmid, Janssen, MoonLake, Novartis, and UCB Pharma; personal fees and grants from Incyte; is a co-copyright holder of HiSQOL; has been a consultant for and received honoraria from AbbVie, Alumis, DermTech, Incyte, Insmid, Janssen, MoonLake Immunotherapeutics, Novartis, and UCB Pharma; and has been a speaker for AbbVie, Janssen, Novartis, and UCB Pharma. EP has been a consultant, advisory board member, speaker for, and received honoraria from Almirall, Janssen-Cilag, GSK, MoonLake Immunotherapeutics, Novartis, and UCB Pharma; and his department received investigator-initiated grant support from AbbVie, Celgene, CHDR, Citryll, Kymera, Janssen-Cilag, and UCB Pharma. JRI receives a stipend as Editor-in-Chief of the British Journal of Dermatology and an authorship honorarium from UpToDate; has been a consultant for AbbVie, Boehringer Ingelheim, ChemoCentryx, Citryll, MoonLake Immunotherapeutics, Novartis, UCB Pharma, and Union Therapeutics, and has served on advisory boards for Insmid, Kymera Therapeutics, and Viela Bio; is a co-copyright holder of the Hidradenitis Suppurativa Quality of Life score (HiSQOL) and hidradenitis suppurativa-Investigator Global Assessment; and his department receives income from copyright of the Dermatology Life Quality Instrument and related instruments. AG has been a consultant and receives honoraria from AbbVie, Aclaris Therapeutics, AnaptysBio, Arista Therapeutics, Bristol Myers Squibb, Boehringer Ingelheim, Incyte, Insmid, Janssen, Novartis, Pfizer, Sonoma Biotherapeutics, UCB Pharma, Union Therapeutics, Ventyx Biosciences, and Viela Biosciences; has received research grants from AbbVie, C3, National Psoriasis Foundation, and UCB Pharma; and is a co-copyright holder of HiSQOL and hidradenitis suppurativa-IGA. ABG reports honoraria as an advisory board member and consultant for Amgen, Almirall, AnaptysBio, Avotres Therapeutics, Boehringer Ingelheim, Bristol Myers Squibb, Dice Therapeutics, Eli Lilly and Company, Janssen, Novartis, Sanofi, UCB Pharma, and Xbiotech; and research or educational grants from AnaptysBio, Bristol Myers Squibb, MoonLake Immunotherapeutics, Novartis, and UCB Pharma (all paid to Mount Sinai School of Medicine). JCJS has acted as a consultant and advisory board member of AbbVie, LEO Pharma, Novartis, Pierre Fabre, Sanofi Genzyme, and Trevi Therapeutics; acted as a speaker for AbbVie, Almirall, Eli Lilly and Company, LEO Pharma, Novartis, Pfizer, Pierre-Fabre, Sanofi, and UCB Pharma; and has acted as an investigator for AbbVie, Amgen, Bristol Myers Squibb, Galapagos, Galderma, Incyte, InflaRX, Janssen, Kliniksa, Kymab Limited, Menlo Therapeutics, Merck, Novartis, Pfizer, Regeneron Pharmaceuticals, Trevi Therapeutics, and UCB Pharma. FGB reports honoraria for participation in advisory boards, in clinical trials, and as a speaker from AbbVie, AbbVie Deutschland, Boehringer Ingelheim, Celltrion, Dr. Wolff, Incyte, Janssen, Mölnlycke, MoonLake Immunotherapeutics, Novartis, and UCB Pharma. E/J-G reports honoraria from Abbott Products Operations, bioMérieux, Brahms, GSK, InflaRX, Sobi and XBiotech; independent educational grants from Abbott Products Operations, AxisShield, bioMérieux, InflaRX, Johnson & Johnson, MSD, Novartis, Sobi, and XBiotech; and funding from the Horizon2020 Marie Skłodowska-Curie International Training Network the European Sepsis Academy (granted to the National and Kapodistrian University of Athens), the Horizon 2020 European Grants ImmunoSep and RISCinCOVID (granted to the Hellenic Institute for the Study of Sepsis) and the Horizon Health grant EPIC-CROWN-2 (granted to the Hellenic Institute for the Study of Sepsis). HF reports honoraria or fees for serving on advisory boards, as a speaker, and as a consultant; and grants as an investigator from AbbVie, Amgen, Boehringer Ingelheim, Bristol Myers Squibb, Daiichi Sankyo, Eisai, Eli Lilly and Company, Janssen, Japan Blood Products Organization, JMEC, Kaken Pharmaceutical, Kyowa Kirin, LEO Pharma, Maruho, Mitsubishi Tanabe,

Nihon Pharmaceutical, Novartis, Otsuka Pharmaceutical, Sanofi, Sato Pharmaceutical, Sun Pharma, Taiho Pharmaceutical, Torii Pharmaceutical, UCB Pharma, and Ushio. RR, PJ, PD, EM, and LP are employees and shareholders of UCB Pharma. CM is a former employee and shareholder of UCB Pharma. MB is an employee of UCB Pharma. CCZ reports grants paid to the institution from AstraZeneca, Boehringer Ingelheim, GSK, InflaRx, Novartis, Relaxera, and UCB Pharma for participation as a clinical and research investigator; consultant honoraria from AccureAcne, Almirall, Boehringer Ingelheim, Incyte, InflaRx, Janssen, L'Oréal, Luvos, NAOS-BIODERMA, Novartis, PPM, Sanofi, UCB Pharma, and Viartis; lecture fees from Almirall, Biogen, Novartis, Sobi, and UCB Pharma; and is the president of the European Hidradenitis Suppurativa Foundation, coordinator of the ALLOCATE Skin group of the European Reference Network for rare skin diseases, and chair of the Acne, Rosacea and Hidradenitis Suppurativa Task Force group of the European Academy of Dermatology and Venereology (EADV); is the editor of the EADV News; and co-copyright holder of the International Hidradenitis Suppurativa Severity Score System on behalf of the European Hidradenitis Suppurativa Foundation.

Data sharing

Underlying data from this manuscript can be requested by qualified researchers 6 months after product approval in the USA and Europe, or global development is discontinued, and 18 months after trial completion. Investigators can request access to anonymised individual patient-level data and redacted trial documents, which can include analysis-ready datasets, study protocol, annotated case-report form, statistical analysis plan, dataset specifications, and clinical study report. Before use of the data, proposals need to be approved by an independent review panel at <https://vivli.org/> and a signed data-sharing agreement will need to be executed. All documents are available in English only, for a prespecified time, typically 12 months, on a password-protected portal.

Acknowledgments

We thank the patients, their caregivers, and the investigators and their teams who took part in these studies. We also acknowledge the UCB clinical programme managers Rob Byerly (UCB Pharma, Morrisville, North Carolina, USA) and Sarah Knowles (UCB Pharma, Slough, UK), Susanne Wiegatz (UCB Pharma, Monheim, Germany) for publication coordination, and Poppy Wilson (Costello Medical, London, UK) and Phoebe Kennedy (Costello Medical, Bristol, UK) for medical writing and editorial assistance. These studies were funded by UCB Pharma. Support for third-party writing assistance for this Article, provided by Phoebe Kennedy, was funded by UCB Pharma in accordance with Good Publication Practice 2022 guidelines (<https://www.ismnp.org/gpp-2022>).

References

- Zouboulis CC, Del Marmol V, Mrowietz U, et al. Hidradenitis suppurativa/acne inversa: criteria for diagnosis, severity assessment, classification and disease evaluation. *Dermatology* 2015; **231**: 184–90.
- Matusiak L. Profound consequences of hidradenitis suppurativa: a review. *Br J Dermatol* 2020; **183**: e171–77.
- Nguyen TV, Damiani G, Orenstein LAV, et al. Hidradenitis suppurativa: an update on epidemiology, phenotypes, diagnosis, pathogenesis, comorbidities and quality of life. *J Eur Acad Dermatol Venereol* 2021; **35**: 50–61.
- Jfri A, Nassim D, O'Brien E, et al. Prevalence of hidradenitis suppurativa: a systematic review and meta-regression analysis. *JAMA Dermatol* 2021; **157**: 924–31.
- Garg A, Neuren E, Cha D, et al. Evaluating patients' unmet needs in hidradenitis suppurativa: results from the Global Survey of Impact and Healthcare Needs (VOICE) project. *J Am Acad Dermatol* 2020; **82**: 366–76.
- Thorlacius L, Cohen AD, Gislason GH, et al. Increased suicide risk in patients with hidradenitis suppurativa. *J Invest Dermatol* 2018; **138**: 52–57.
- Matusiak L, Bieniek A, Szepietowski JC. Hidradenitis suppurativa markedly decreases quality of life and professional activity. *J Am Acad Dermatol* 2010; **62**: 706–08.
- Markota Čagalj A, Marinović B, Bukvić Mokos Z. New and emerging targeted therapies for hidradenitis suppurativa. *Int J Mol Sci* 2022; **23**: 1.
- Kimball AB, Jemec GBE, Alavi A, et al. Secukinumab in moderate-to-severe hidradenitis suppurativa (SUNSHINE and SUNRISE): week 16 and week 52 results of two identical, multicentre, randomised, placebo-controlled, double-blind phase 3 trials. *Lancet* 2023; **401**: 747–61.
- Kimball AB, Okun MM, Williams DA, et al. Two phase 3 trials of adalimumab for hidradenitis suppurativa. *N Engl J Med* 2016; **375**: 422–34.
- Zouboulis CC, Nogueira da Costa A, Makrantonaki E, et al. Alterations in innate immunity and epithelial cell differentiation are the molecular pillars of hidradenitis suppurativa. *J Eur Acad Dermatol Venereol* 2020; **34**: 846–61.
- Frew JW, Hawkes JE, Krueger JG. A systematic review and critical evaluation of inflammatory cytokine associations in hidradenitis suppurativa. *F1000Res* 2018; **7**: 1930.
- Skelton A, Gopee H, Okoye R, et al. IL-17A and IL-17F in the pathogenesis of hidradenitis suppurativa: insights from early and chronic lesions. *J Invest Dermatol* 2023; **143**: S87.
- Cole S, Manghera A, Burns L, et al. Differential regulation of IL-17A and IL-17F via STAT5 contributes to psoriatic disease. *J Allergy Clin Immunol* 2023; **152**: 783–98.
- Glatt S, Baeten D, Baker T, et al. Dual IL-17A and IL-17F neutralisation by bimekizumab in psoriatic arthritis: evidence from preclinical experiments and a randomised placebo-controlled clinical trial that IL-17F contributes to human chronic tissue inflammation. *Ann Rheum Dis* 2018; **77**: 523–32.
- Adams R, Maroof A, Baker T, et al. Bimekizumab, a novel humanized IgG1 antibody that neutralizes both IL-17A and IL-17F. *Front Immunol* 2020; **11**: 1894.
- Reich K, Warren RB, Lebwohl M, et al. Bimekizumab versus secukinumab in plaque psoriasis. *N Engl J Med* 2021; **385**: 142–52.
- Glatt S, Jemec GBE, Forman S, et al. Efficacy and safety of bimekizumab in moderate to severe hidradenitis suppurativa: a phase 2, double-blind, placebo-controlled randomized clinical trial. *JAMA Dermatol* 2021; **157**: 1279–88.
- Ingram JR, Ciaravino V, Roller R, et al. Electronic patient-reported outcomes in hidradenitis suppurativa: content validity and usability of the electronic hidradenitis suppurativa symptom daily diary, hidradenitis suppurativa symptom questionnaire and hidradenitis suppurativa quality of life questionnaire. *Dermatology* 2023; **240**: 65–76.
- Kimball AB, Sobell JM, Zouboulis CC, et al. HiSCR (hidradenitis suppurativa clinical response): a novel clinical endpoint to evaluate therapeutic outcomes in patients with hidradenitis suppurativa from the placebo-controlled portion of a phase 2 adalimumab study. *J Eur Acad Dermatol Venereol* 2016; **30**: 989–94.
- Basra MK, Salek MS, Camilleri L, et al. Determining the minimal clinically important difference and responsiveness of the Dermatology Life Quality Index (DLQI): further data. *Dermatology* 2015; **230**: 27–33.
- Merola JF, Landewé R, McInnes IB, et al. Bimekizumab in patients with active psoriatic arthritis and previous inadequate response or intolerance to tumour necrosis factor- α inhibitors: a randomised, double-blind, placebo-controlled, phase 3 trial (BE COMPLETE). *Lancet* 2023; **401**: 38–48.
- van der Heijde D, Deodhar A, Baraliakos X, et al. Efficacy and safety of bimekizumab in axial spondyloarthritis: results of two parallel phase 3 randomised controlled trials. *Ann Rheum Dis* 2023; **82**: 515–26.
- McInnes IB, Asahina A, Coates LC, et al. Bimekizumab in patients with psoriatic arthritis, naive to biologic treatment: a randomised, double-blind, placebo-controlled, phase 3 trial (BE OPTIMAL). *Lancet* 2023; **401**: 25–37.
- Zouboulis CC, Gulliver W, Ingram J, et al. Endpoints of clinical trials for hidradenitis suppurativa: proceedings of a round-table session. *Exp Dermatol* 2020; **29** (suppl 1): 67–72.
- Moltrasio C, Tricarico PM, Romagnuolo M, et al. Hidradenitis suppurativa: a perspective on genetic factors involved in the disease. *Biomedicine* 2022; **10**: 2039.
- Zouboulis CC. Pain Index: a new prospective hidradenitis suppurativa patient-reported outcome measure instrument. *Br J Dermatol* 2021; **184**: 1203–04.
- Zouboulis CC, Tzellos T, Kyrgidis A, et al. Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHSS4), a novel dynamic scoring system to assess hidradenitis suppurativa severity. *Br J Dermatol* 2017; **177**: 1401–09.

-
- 29 Thorlacius L, Esmann S, Miller I, et al. Development of HiSQOL: a hidradenitis suppurativa-specific quality of life instrument. *Skin Appendage Disord* 2019; **5**: 221–29.
- 30 Zouboulis CC, Desai N, Emtestam L, et al. European S1 guideline for the treatment of hidradenitis suppurativa/acne inversa. *J Eur Acad Dermatol Venereol* 2015; **29**: 619–44.
- 31 Davidson L, van den Reek J, Bruno M, et al. Risk of candidiasis associated with interleukin-17 inhibitors: a real-world observational study of multiple independent sources. *Lancet Reg Health Eur* 2022; **13**: 100266.
- 32 Petitpain N, D'Amico F, Yelehe-Okouma M, et al. IL-17 inhibitors and inflammatory bowel diseases: a postmarketing study in vigibase. *Clin Pharmacol Ther* 2021; **110**: 159–68.
- 33 Ruggiero A, Martora F, Picone V, et al. Paradoxical hidradenitis suppurativa during biologic therapy, an emerging challenge: a systematic review. *Biomedicines* 2022; **10**: 455.

THE LANCET

Supplementary appendix 1

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Kimball AB, Jemec GBE, Sayed CJ, et al. Efficacy and safety of bimekizumab in patients with moderate-to-severe hidradenitis suppurativa (BE HEARD I and BE HEARD II): two 48-week, randomised, double-blind, placebo-controlled, multicentre phase 3 trials. *Lancet* 2024; published online May 22. [https://doi.org/10.1016/S0140-6736\(24\)00101-6](https://doi.org/10.1016/S0140-6736(24)00101-6).

1 SUPPLEMENTARY APPENDIX 1

2 This appendix has been provided by the authors to give readers additional information about their
3 work.

4 Title

5 Bimekizumab efficacy and safety in patients with moderate to severe hidradenitis suppurativa: two 48-
6 week, randomised, double-blind, placebo-controlled, multicentre phase 3 trials (BE HEARD I and II)

7 Authors

8 Alexa B. Kimball,* Gregor B. E. Jemec,* Christopher Sayed, Joslyn S. Kirby, Errol Prens, John R.
9 Ingram, Amit Garg, Alice B. Gottlieb, Jacek C. Szepietowski, Falk G. Bechara, Evangelos J.
10 Giamarellos-Bourboulis, Hideki Fujita, Robert Roller, Paulatsya Joshi, Pratiksha Dokhe, Edward
11 Muller, Luke Peterson, Cynthia Madden, Muhammad Bari, Christos C. Zouboulis

12 *Joint Primary Authorship

13 Table of Contents

14		
15	SUPPLEMENTARY APPENDIX 1	1
16	BE HEARD I: List of principal investigators and study sites	3
17	BE HEARD II: List of principal investigators and study sites	6
18	Supplementary Methods	10
19	Full list of inclusion criteria.....	10
20	Full list of exclusion criteria.....	10
21	Independent DMC and subject matter expert adjudication committees	13
22	Clinical definitions	14
23	Statistical analyses.....	14
24	Exploratory endpoints in BE HEARD I and BE HEARD II	15
25	Supplementary Results	17
26	Primary efficacy outcome in BE HEARD I and BE HEARD II at week 16 using OC data.....	17
27	Supplementary Tables	18
28	Table S1. Number of patients with HiSCR50 data observed or imputed at week 16 in BE HEARD I and	
29	BE HEARD II	18
30	Table S2. HiSCR90 and HiSCR100 responses at week 16 and week 48	19
31	Table S3. AN and DT count at baseline, week 16 and week 48.....	20
32	Table S4. TEAEs with BKZ treatment to week 16 in BE HEARD I and BE HEARD II	21
33	Table S5. TEAEs by severity grade category to week 16 in BE HEARD I and BE HEARD II	22
34	Table S6. Severe grade TEAEs to week 16 in BE HEARD I and II	22
35	Table S7. TEAEs by treatment regimen to week 48 in BE HEARD I and BE HEARD II	23
36	Supplementary Figures	25
37	Figure S1. Study design.....	25
38	Figure S2. Sequence of testing	26
39	Figure S3. Missing data handling	28
40	Figure S4. HiSCR50 and HiSCR75 responses over time to week 48.....	29
41	Figure S5. HiSCR90 responses over time to week 48.....	30

42	Figure S6. HiSCR100 responses over time to week 48.....	31
43	Figure S7. Change from baseline in AN count to week 48	32
44	Figure S8. Change from baseline in DT count to week 48.....	33
45	Figure S9. HSSDD Worst Skin Pain Response over time to week 16.....	34
46		
47		

48 **BE HEARD I: List of principal investigators and study sites**

Principal investigator	Study site	Patients recruited (N)
Evangelos Giamarellos-Bourboulis	4 th Department of Internal Medicine, ATTIKON University Hospital, 1 Rimini Str., Athens, Attiki, Greece, 12462	40
Seth Forman	15416 N. Florida Ave., Tampa, Florida, United States, 33613	32
Lindsay Ackerman	1331 N. 7 th Street, Suite 250, Phoenix, Arizona, United States, 85006	23
Christopher Sayed	410 Market Street, Chapel Hill, North Carolina, United States, 27516	16
Mark Lebwohl	5 East 98 th Street, Dermatology Clinical Trials – 5 th Floor, New York, New York, United States, 10029	15
Giuseppe Micali	78, Via S. Sofia, Catania, Catania, Italy, 95123	14
Falk Bechara	Gudrunstr. 56, Klinikum der Ruhr-Universität Bochum, Bochum, Nordrhein-Westfalen, Germany, 44791	12
Anne-Benedicte Duval Modeste	Clinique Dermatologique – Hôpital Charles Nicolle, 1, Rue de Germont, Rouen, Seine-Maritime, France, 76031	10
Maurizio Podda	Klinikum Darmstadt, Hautklinik, Darmstadt, Hessen, Germany, 64283	10
Errol Prens	Dr. Molewaterplein 40, Rotterdam, Zuid-Holland, Netherlands, 3015 GD	10
Ben Bogucki	150 Taylor Station Road, Suite 100, Columbus, Ohio, United States, 43213	10
Lorne Albrecht	13737 96 Avenue, Suite 603, Surrey, British Columbia, Canada, V3V 0C6	9
Georgios Kokolakis	Charita Platz 1, Berlin, Berlin, Germany, 10117	9
Lara Ferrándiz Pulido	Hospital Universitario Virgen de la Macarena, Avda. Doctor Fedriani, 3. Dep. Dermatología, Edificio Policlínico, 2 Planta, Sevilla, Andalucía, Spain, 41009	9
Angeles Florez Menéndez	Complejo Hospitalario, Universitario de Pontevedra, Servicio de Dermatología, C/ Simon Bolivar, 6, Pontevedra, Pontevedra, Spain, 36001	9
Howard Sofen	Dermatology Research Associates, 8930 S. Sepulveda Blvd., Suite 114, Los Angeles, California, United States, 90045	9
David Greenstein	138 Conant Street, First Floor, Beverly, Massachusetts, United States, 01915	9
James Del Rosso	9080 West Post Road, Suite 100, Las Vegas, Nevada, United States, 89148	9
Kirsten Walker	39 23 rd Street East, Suite 1010, Saskatoon, Saskatchewan, Canada, S7K 0H6	8
Maryam Shayesteh-Alam	5 Quarry Ridge Road, Suite 105, Barrie, Ontario, Canada, L4M 7G1	8
Anne-Claire Fougousse	69 Avenue de Paris, Service de Dermatologie, Saint Mande Cedex, Val-de-Marne, France, 94163	8
Nathalie Beneton-Benhard	194, Avenue Rubillard, Service Dermatologie, Le Mans, Sarthe, France, 72037	8
Jerry Bagel	59 One Mile Rd., 2 nd and 3 rd Floor, East Windsor, New Jersey, United States, 08520	8
Angela Moore	711 East Lamar Blvd., Suite 100, 200, Arlington, Texas, United States, 76011	8
Petra Staubach-Renz	Langenbeckstr. 1, Mainz, Rheinland-Pfalz, Germany, 55131	7
Alexa Kimball	330 Brookline Avenue, Boston, Massachusetts, United States, 2215	7
Hollis Clark	205 South 6 th Street, Murray, Kentucky, United States, 42071	7
Mads Rasmussen	P.P Oerumsgade 11, Aarhus, Denmark, 8000	6
Thierry Boye	2 Boulevard Saint Anne, Service de Dermatologie, Toulon Cedex 9, Var, France, 83800	6
Antonio Costanzo	56, Via Manzoni, Rozzano, Milano, Italy, 20089	6

Boni Elewski	2000 6 th Avenue South, 3 rd Floor, Birmingham, Alabama, United States, 35233	6
Michael Gold	2000 Richard Jones Road, Suite 223, Nashville, Tennessee, United States, 37215	6
Walter Nahm	7695 Cardinal Court, Suite 210, San Diego, California, United States, 92123	6
Rodney Sinclair	Sinclair Dermatology Level 3, 2 Wellington Pde, East Melbourne, Victoria, Australia, 3002	5
Shireen Sidhu	562 Lower North East Road, North Eastern Health Specialists Campbelltown, South Australia, Australia, 5074	5
Deirdre Murrell	17 Kensington Street, The Church, Kogarah, New South Wales, Australia, 2217	5
Mani Raman	312 Highway 7 East, Richmond Hill, Ontario, Canada, L4B 1A5	5
Nina Magnolo	Klinik for Hautkrankheiten, Von-Esmarch-Str. 58, Münster, Nordrhein-Westfalen, Germany, 48149	5
Sascha Gerdes	Schwanenweg 20, Kiel, Schleswig-Holstein, Germany, 24105	5
Elisavet Lazaridou	Nea Periferiaki Odos Thessalonikis – Nea Efkarpi, Thessaloniki, Kentriki Makedonia, Greece, 56403	5
Gregorio Carretero Hernández	Hospital Universitario de Gran Canaria Doctor Negrín, Pl. Barranco la Ballena, s/n, Servicio de Dermatología, 6 Planta, Las Palmas de Gran Canaria, Las Palmas, Spain, 35019	5
Serhat Inaloz	Gaziantep Universitesi Sahinbey Arastirma ve Uygulama Hastanesi, Deri ve Zuhrevi Hastaliklari Bilimdali, Sehitkamil, Gaziantep, Turkey, 27310	5
Debra Breneman	222 Piedmont Avenue, Suite 5300, Cincinnati, Ohio, United States, 45219	5
Sinan Dongan	Bozyaka Training and Research Hospital, Izmir Buca Seyfi Demirsoy Public Hospital Bahar, Saim Cikrikci Cd. No: 59, Izmir, 35170, Turkey	5
Alison Fischer	7302 S. Yale Ave., Tulsa, Oklahoma, United States, 74136	5
Peter Foley	80 Drummond Street Level 1, Skin and Cancer Foundation Inc. Carlton, Victoria, Australia, 3053	4
Thierry Passeron	151, Route de Saint Antoine Ginestière, Service de Dermatologie, Nice Cidex 3, Alpes-Maritimes, France, 06202	4
Barbara Horváth	Hanzeplein 1, Groningen, Groningen, Netherlands, 9713 GZ	4
Sebastiaan Prens	Molengracht 21, Breda, Zuid-Holland, Netherlands, 4818 CK	4
Julio Bassas Vila	Carretera de Canyet s/n, Badalona, Barcelona, Spain, 08916	4
Burhan Engin	Cerrahpasa Tip Fakultesi Yerleskesi Kocamustafapasa Cd., 53, Cerrahpasa 34098 Fatih, Istanbul, Istanbul, Turkey, 34098	4
Kathleen Gordon	1524 Atwood Avenue, Suite 330, Johnston, Rhode Island, United States, 02919	4
Mark Lee	1973 NW Loop 410, Suite 106, San Antonio, Texas, United States, 78213	4
Tory Sullivan	16100 NE 16 th Ave., Suite A, North Miami Beach, Florida, United States, 33162–4708	4
Mack Rachal	1110 West Peachtree Street, NW, Suite 1040, Atlanta, Georgia, United States, 30309	4
Lynda Spelman	Suite 18, Level 1, 250 Ipswich Rd, Woolloongabba, Queensland, Australia, 4102	3
An Van Laethem	Herestraat 49, Leuven, Vlaams Brabant, Belgium, 3000	3
Rafael Botella Estrada	Hospital Universitari i Politecnic La Fe, Avda. Fernando Abril Martorell, 106, Valencia, Valencia, Spain, 46026	3
Erkan Alpsoy	Akdeniz Universitesi Tip Fakultesi, Dermatoloji Bilim Dadi, Konyaalti, Antalya, Turkey, 07070	3
Arjen Nikkels	CHU Liege – Domaine Universitaire du Sart Tilman B35, Departement de Dermatologie, Liege, Belgium, 4000	2

Pierre-Dominique Ghislain	Cliniques Universitaires Saint Luc, Avenue Hippocrate 10, Brussels, Brussels Capital Region, Belgium, 1200	2
Dagmar Wilsmann-Theis	Venusberg-Campus 1, Bonn, Nordrhein-Westfalen, Germany, 53127	2
Ariela Hafner	6 Weizmann Street, Tel-Aviv, Israel, 6423906	2
Ketty Peris	Largo Agostino Gemelli 8, 9 Piano Ala E., Stanza E925, Roma, Roma, Italy, 00168	2
Paolo Dapavo	23, Via Cherasco, Presidio S. Lazzaro A.O.U. Citta della Salute e della Scienza, Torino, Torino, Italy, 10126	2
Marit Saunes	Mauritz Hansens, Gate 3, Trondheim, Norway, 7030	2
Robert Hunger	Inselspital Bern, Universitätsklinik für Dermatologie, Freiburgstrasse 34, Bern, Bern, Switzerland, 3010	2
Emmanuel Laffitte	Rue Gabrielle-Perret-Gentil 4, Geneve, Geneve, Switzerland, 1211	2
Joslyn Kirby	500 University Dr., UPC II, Room Rm 2010, Hershey, Pennsylvania, United States, 17033	2
Carlos Wambier	593 Eddy Street, Rhode Island Hospital, APC-10, Providence, Rhode Island, United States, 02903	2
Joel Schlessinger	2802 Oak View Drive, Omaha, Nebraska, United States, 68144	2
Véronique Del Marmol	Route de Lennik 808, Brussels, Brussels Capital Region, Belgium, 1070	1
Guillaume Chaby	2 Place Victor Pauchet, Service de Dermatologie et Venerologie, Amiens, Picardie, France, 80080	1
Juliette Delaunay	4 Rue Larrey, Service de Dermatologie, Angers, France, 49100	1
Franziska Rueff	Frauenlobstrae 9–11, Ludwig-Maximilians-Universität München, Klinik und Poliklinik für Dermatologie und Allergologie, München, Bayern, Germany, 80337	1
Uwe Wollina	Friedrichstr. 4, Dermatologie und Allergologie, Dresden, Sachsen, Germany, 01067	1
Ziad Khamaysi	8 Haalia St., Bat Galim, Haifa, Israel, 3109601	1
James Solomon	725 West Granada Blvd., Suite 44, Ormond Beach, Florida, United States, 32174	1
Martin Menter	3900 Junius St., Suite 125, Dallas, Texas, United States, 75246	1
Ross Campbell	1180 Resurgence Dr., Watkinsville, Georgia, United States, 30677	1
Wayne Gulliver	187 LeMarchant Road, Suite 007, Saint-John's, Newfoundland and Labrador, Canada, A1C 2H5	0
Evelyn Gaffal	Universitätsklinikum Magdeburg, Universitshautklinik, Haus 14, Magdeburg, Sachsen-Anhalt, Germany, 39120	0
Dimitrios Rigopoulos	Rimini St. 1, Athens, Greece, 12462	0
Luca Bianchi	81, Via Oxford, Roma, Roma, Italy, 00133	0
Basak Yalici Armagan	Hacettepe Üniversitesi Tıp Fakültesi, Dermatoloji Anabilim Dalı, Ankara, Ankara, Turkey, 06230	0
Peter Jenkin	1730 Minor Avenue, Suite 1000, Seattle, Washington, United States, 98101	0

49 Recruited patients included all those enrolled and included in the randomised set. Investigators listed
50 here with '0' had patients who were screened but did not proceed to be randomised.

51 **BE HEARD II: List of principal investigators and study sites**

Principal investigator	Study site	Patients recruited (N)
Jacek Szepietowski	CITYCLINIC PRZYCHODNIA LEKARSKO-PSYCHOLOGICZNA MATUSIAK SPÓŁ, ul. Sliczna 13, Wrocław, Poland, 50-566	40
Wojciech Baran	WroMedica I. Bielicka, A. Strzalkowska s.c., ul. Mickiewicza 91, Wrocław, Poland, 51-685	31
Kamelia Vekovska	MC Medconsult Pleven – Dermatology, 18, Sv. Kiril I Metodi, Pleven, Plevna, Bulgaria, 5800	31
Irena Walecka - Herniczek	Centralnego Szpitala Klinicznego MSWiA, ul. Woloska 137, Warszawa, Poland, 02-507	21
Monika Arenbergerova	Fakultni nemocnice v Motole, V Uvalu 84, Praha 5, Praha 5, Czech Republic, 150 06	16
Carle Paul	CHU de Toulouse – Hopital Larrey – Dermatologie, 24, Chemin de Pouvoirville – TSA 30030, Occitanie, Toulouse, France, 31400	15
Jolanta Weglowska	DermMedica Sp. z o.o., ul. Zakrzowska 19 A, Wrocław, Poland, 51-318	13
Ziad Regui	Polyclinique De Courlancy, 38 Rue De Courlancy, Avenue Du General Koenig, Marne, Reims, France, 51100	12
Andreas Pinter	Goethe University Hospital, Theodor Stern-Kai 7, Frankfurt/Main, Hessen, Germany, 60590	12
Adam Reich	Kliniczny Szpital Wojew. nr 1 im. F.Chopina w Rzeszowie, ul. Szopena 2, Rzeszów, Poland, 35-055	12
Matthew Miller	The Indiana Clinical Trials Center – Research, 900 Mohawk Street, Plainfield, Indiana, United States, 46168	12
Mariya Ratkova	Medical Center “Hera” EOOD, 20, Klisura St., Sofia, Sofia-Grad, Bulgaria, 1510	11
Jaggi Rao	Alberta DermaSurgery Centre, Suite 300, 7609–109 Street, NW, Edmonton, Alberta, Canada, T6G 1C3	11
George Schmieder	Park Avenue Dermatology, 906 Park Avenue, Orange Park, Florida, United States, 32073	10
Scott Guenther	Georgia Skin and Cancer Clinic, 900 Mohawk Street, Suite E, Savannah, Georgia, United States, 31419	10
Marni Wiseman	Wiseman Dermatology Research Inc., 1170 Taylor Avenue, Unit 6, Winnipeg, Manitoba, Canada, R3M 3Z4	8
Florian Schenck	Hautarzte Zentrum Hannover, Osterstr. 24, Hannover, Niedersachsen, Germany, 30159	8
Robert Snyder	Riverchase Dermatology and Cosmetic Surgery, 603 North Flamingo Rd. Suite 350, Pembroke Pines, Florida, United States, 33028	8
Jamie Weisman	Advanced Medical Research PC, 5730 Glenridge Drive NE, Suite T-100, Sandy Springs, Georgia, United States, 30329	8
Tarannum Jaleel	Wright State Physicians, Inc725 University Blvd., Fairborn, Ohio, United States, 45324	8
Elmira Pavlova	Medical Center Robert Koch, 16V Haga Str., Sofia, Sofia-Grad, Bulgaria, 1407	7
Alena Machovcova	FN Kralovske Vinohrady, Srobarova 50, Praha 10, Praha 10, Czech Republic, 100 34	7
Andrea Szegedi	Debreceni Egyetem Klinikai Központ Nagyerdei Campus Gyermekgyógyászati Klinika, Nagyerdei Körút 98, Hajdú-Bihar, Debrecen, Hungary, 4032	7
Joanna Narbutt	Dermoklinika Centrum Medyczne s.c. M.Kierstan, J.Narbutt, A.Lesiak, Al. Kosciuszki 93, Łódź, Poland, 90-436	7
Wendy McFald	Clarkston Skin Research, 5701 Bow Pointe Drive, Suite 217, Clarkston, Michigan, United States, 48346	7
Hadar Lev-Tov	Clinical Studies Group, 3095 Old Conejo Rd, Suite 110, Henderson, Nevada, United States, 89052	7

Grisha Mateev	Diagnostic Consultative Center Fokus-5 Medical Establishment for Outpatient Care EOOD, 15 Hristo Stanchev Str., Sofia, Sofia, Bulgaria, 1463	6
Melinda Gooderham	SKiN Health – Clinic, 1176 Division Street, Cobourg, Ontario, Canada, K9A 0Z4	6
Pierre-Andre Becherel	Hopital Prive D'Antony, 1 Rue Velpeau, Hauts-de-Seine, Antony, France, 92160	6
Sidney Smith	Duke University Medical Center, 40 Medicine Circle, Durham, North Carolina, United States, 27710	6
Kim Papp	K. Papp Clinical Research Inc., 135 Union Street East, Waterloo, Ontario, Canada, N2J 1C4	5
Wei Loo	DermEffects, 1560 Hyde Park Road, London, Ontario, Canada, N6H 5L5	5
Marie Aleth Richard Lallemand	Aphm – Hopital De La Timone – Neurology, 254 Rue Saint-Pierre, Bouches-du-Rhône, Marseille, France, 13005	5
Natalia Kirsten	Universitätsklinikum Hamburg-Eppendorf, Martinistr. 52, Hamburg, Hamburg, Germany, 20246	5
Jeffrey Travers	The Dermatology Group, PC, 60 Pompton Ave., Verona, New Jersey, United States, 07044	5
Evgeniya Hristakieva	UMHAT "Prof. Kirkovich", 2, General Stoilev St., Stara Zagora, Stara Zagora, Bulgaria, 6003	4
Olivier Cogrel	CHU de Bordeaux Hopital Saint-Andre, 1 Rue Jean Burguet, Gironde, Bordeaux Cedex, France, 33075	4
Diamant Thaci	UK-SH – Lubeck, Ratzeburger Allee 160, Haus 50, Lübeck, Schleswig-Holstein, Germany, 23538	4
Irene Fuertes	Hospital Clinic de Barcelona, Calle Villarroel 170, Barcelona, Barcelona, Spain, 08036	4
Antonio Martorell Calatayud	Hospital de Manises, Avda. Generalitat Valenciana, 50, Manises, Valencia, Spain, 46940	4
Milan Anadkat	Washington University – School of Medicine, 660 South Euclid Ave, Campus Box 8123, Saint Louis, Missouri, United States, 63110	4
Abel Jarell	ActivMed Practices & Research, LLC., 110 Corporate Drive, Suite 2, Portsmouth, New Hampshire, United States, 03801	4
Vivian Laquer	First OC Dermatology, 17271 Brookhurst Street, Fountain Valley, California, United States, 92708	4
Navid Ezra	California Dermatology Institute, 3095 Old Conejo Rd, Suite 200, Thousand Oaks, California, United States, 91320	4
Hidetoshi Takahashi	Takagi Dermatological Clinic – Dermatology, Nishi3-Minami4-16, Hokkaido, Obihiro, Japan, 080-0013	4
Philippe Celerier	La Rochelle – Re – Aunis, Rue du Docteur Schweitzer, Marselle, La Rochelle, France, 17019	3
Hugues Barthelemy	Centre Hospitalier d'Auxerre, 2 Boulevard De Verdun, Bourgogne, Auxerre, France, 89011	3
Jean-Luc Perrot	CHU Saint-Etienne – Dermatologie et Venereologie, Avenue Albert Raimond, Loire, Saint-Étienne, France, 42277	3
Eva Vilarrasa Rull	Hospital de La Santa Creu i Sant Pau, Hospital de la Santa Creu i Sant Pau. Dermatologia, Sant Antoni Maria Claret 167, Barcelona, Barcelona, Spain, 08025	3
Alejandro Molina-Leyva	Hospital Universitario Virgen de las Nieves C/ Doctor Azpitarte, 4, Granada, Andalucía, Spain, 18012	3
Rita Pichardo-Geisinger	Wake Forest Baptist Medical Center, 4618 Country Club Road, Winston Salem, North Carolina, United States, 27104	3
Robert Noss	University of Miami Hospital, 1321 NW 14th Street, West Bldg., Room 505, Miami, Florida, United States, 33125	3
Keiko Hashikawa	Kurume University Hospital – Neurology, 67 Asahi-machi, Fukuoka, Kurume, Japan, 830-0011	3
Toshiya Takahashi	Tohoku University Hospital, 1-1 Seiryomachi, Aoba-ku, Miyagi, Sendai, Japan, 980-8574	3

Takuya Miyagawa	The University of Tokyo Hospital, 7-3-1 Hongo, Tokyo, Bunkyo-ku, Japan, 113-8655	3
Johannes Kern	Royal Melbourne Hospital, Reserve Road, Parkville, Victoria, Australia, 3050	2
Gayle Fischer	Royal North Shore Hospital, NSCC, Level 1, RNSH, Reserve Road, St. Leonards, New South Wales, Australia, 2065	2
Ivan Botev	Ambulatory for Specialized Medical Help, Persenk Str., Bl. 19, Sofia, Sofia, Bulgaria, 1407	2
Pick Woo	Northampton General Hospital NHS Trust, Research and Development Unit, Northampton, United Kingdom, NN1 5BD	2
Christos Zouboulis	Städtisches Klinikum Dessau, Auenweg 38, Dessau, Sachsen-Anhalt, Germany, 06847	2
Sylke Schneider-Burrus	Havelklinik GmbH & Co. KG – Zentrum für Dermatochirurgie, Gatower Straue 191, Berlin, Germany, 13595	2
Michael Ziv	Ha'emek Medical Center, 21 Rabin Road, Afula, Ha Zafon, Israel, 18101	2
Stefan Beissert	Universitätsklinikum Carl Gustav Carus, Fetscherstrasse 74, Dresden, Sachsen, Germany, 01307	2
Francisco García Martínez	Clinica Universidad de Navarra, Marquesado de Sta Marta 1, Madrid, Madrid, Spain, 28027	2
Lara Wine Lee	Medical University Of South Carolina, 135 Rutledge Ave, RT 323, Johns Island, South Carolina, United States, 29425	2
Jessica Kaffenberger	Ohio State Medical Center Dermatology East, 540 Officecenter Place, Suite 240, Columbus, Ohio, United States, 43230	2
Edward Lain	Austin Institute for Clinical Research, Inc. – Dermatology, 540 Officecenter Place, Suite 1101, Pflugerville, Texas, United States, 78660	2
Stephanie Mehlis	NorthShore University HealthSystem, 9933 Woods Drive, Suite 100, Skokie, Illinois, United States, 60077	2
Koremasa Hayama	Nihon University Itabashi Hospital, 30-1, Oyaguchi, Kami-cho, Tokyo, Itabashi-ku, Japan, 173-8610	2
Shinichi Imafuku	Fukuoka University Hospital, 7-45-1, Nanakuma, Jonan-ku, Fukuoka, Japan, 814-0180	2
Kentaro Hayashi	Hospital, University of the Ryukyus – Dermatology, 207 Uehara, Nishihara-cho, Okinawa, Nakagami-gun, Japan, 903-0215	2
Masatoshi Abe	Kojinkai Sapporo Skin Clinic, 1-1 Nishi2 Minami3 Chuo-ku, H&B Plaza Building 5F, Hokkaido, Sapporo, Japan, 060-0063	2
Emily Altman	Palmetto Clinical Trial Services, LLC, 37 Brendan Way, Greenville, South Carolina, United States, 29615	2
Akimichi Morita	Nagoya City University Hospital, 1 Kawasumi Mizuhocho Mizuho-ku, Nagoya, Aiti, Japan, 467-8602	2
Pablo Fernandez-Penas	Westmead Hospital, Department of Dermatology, D5a, CD Block Hawkesbury Rd., Westmead, New South Wales, Australia, 2145	1
Kamal Krishna Ohson	Karma Clinical Trials Inc., 63 Airport Road, St. John's, Newfoundland and Labrador, Canada, A1A 4Y3	1
Axel Villani	CHU Hopital Edouard Herriot, 5, Place d'Arsonval, Rhone, Lyon, France, 69437	1
Jens Rossbacher	Hautzentrum Friedrichshain, Frankfurter Allee 100, Berlin, Germany, 10247	1
Brian Kirby	Elm Park, 196 Merrion Road, Dublin 4, Dublin, Ireland, D04 Y8V0	1
Jorge Romani	Hospital Universitari Parc Tauli, C/ Parc Taulí, s/n, Sabadell, Barcelona, Spain, 08208	1
Wedad Abdelrahman	Guy's Hospital, Great Maze Pond, London, United Kingdom, SE1 9RT	1
Philip Hampton	The Newcastle upon Tyne Hospitals NHS Foundation Trust, Westgate Road, Newcastle Upon Tyne, United Kingdom, NE4 6BE	1
Kave Shams	University of Leeds, Wellcome Trust Brenner Building, Leeds, United Kingdom, LS9 7TF	1
Ichiro Kurokawa	Meiwa Hospital – Dermatology, 4-31 Agenaruo-cho, Hyogo, Nishinomiya, Japan, 663-8186	1

Jun Asai	University Hospital Kyoto Prefectural University of Medicine, 465 Kajii-cho, Kawaramachi-hirokoji, Kamigyo-ku, Kyoto, Japan, 602-8566	1
Takuro Kanekura	Kagoshima University Hospital – Neurosurgery, 8-35-1 Sakuragaoka, Kagoshima, Kagoshima, Japan, 890-8520	1
Daisuke Tsuruta	Osaka Metropolitan University Hospital – Rheumatology/Orthopedics, 1-5-7 Asahimachi, Abeno-ku, Osaka, Japan, 545-8586	1
John Ingram	University Hospital of Wales, Cardiff, United Kingdom, CF14 4XW	0
Debra Grayman	Accel Research Sites – PureSkin Dermatology & Aesthetics, 7932 West Sand Lake Road, Orlando, Florida, United States, 38219	0
Lawrence Parish	Paddington Testing Co. Inc., 1845 Walnut Street, Suite 1650, Philadelphia, Pennsylvania, United States, 19103	0

52 Recruited patients included all those enrolled and included in the randomised set. Investigators listed
53 here with '0' had patients who were screened but did not proceed to be randomised.

54 **Supplementary Methods**

55 **Full list of inclusion criteria**

56 Patients are eligible to be included in the study only if all of the following criteria apply:

57 Age

58 1. Patient must be at least 18 years of age, at the time of signing the informed consent. If a study
59 patient is under the local age of consent and is at least 18 years of age, written informed
60 consent will be obtained from both the study patient and the legal representative.

61 Type of patient and disease characteristics

62 2. Study patients must have a diagnosis of HS based on clinical history and physical examination
63 for at least six months prior to the baseline visit; diagnosis must be verifiable through medical
64 notes and documentation.

65 3. Study patient must have HS lesions present in at least two distinct anatomic areas (e.g., left
66 and right axilla), one of which must be at least Hurley Stage II or Hurley Stage III at both the
67 screening and baseline visits.

68 4. Study patient must have moderate to severe HS defined as a total of ≥ 5 inflammatory lesions
69 (i.e., number of abscesses plus number of inflammatory nodules) at both the screening and
70 baseline visits.

71 5. Study patient must have had a history of inadequate response to a course of a systemic
72 antibiotics for treatment of HS at the screening visit as assessed by the investigator through
73 study patient interview and review of medical history; inadequate response must be verifiable
74 through medical notes and documentation. Study patients who meet any of the following are
75 NOT automatically excluded from the study:

76 a. Demonstrated intolerance to (or during therapy became intolerant to) systemic
77 antibiotics

78 b. Had a contraindication to systemic antibiotics

79 c. Responded to course(s) of systemic antibiotic(s) and subsequently exhibited
80 recurrence after discontinuation of the antibiotic

81 Sex

82 6. Males and females may be study patients.

83 a. A female study patient is eligible to participate if she is not pregnant (see Study
84 Protocol [Section 10.4], not breastfeeding, and at least one of the following
85 conditions applies: Not a woman of childbearing potential (WOCBP) as defined in
86 Appendix 4 of the Study Protocol OR a WOCBP who agrees to follow the
87 contraceptive guidance during the treatment period and for at least 20 weeks after the
88 last dose of IMP.

89 Informed consent

90 7. Study patient was capable of giving signed informed consent as described in Appendix 1 of
91 the Study Protocol which includes compliance with the requirements and restrictions listed in
92 the informed consent form (ICF) and in the protocol.

93 **Full list of exclusion criteria**

94 Study patients are excluded from the study if any of the following criteria apply:

95 Medical conditions

96 1. Study patient has any medical or psychiatric condition that, in the opinion of the investigator,
97 could jeopardise or would compromise the study patient's ability to participate in this study as
98 determined by the investigator based on protocol-required assessments.

99 Hidradenitis suppurativa, skin-specific, and other inflammatory disease

100 2. Study patient has a DT count of >20 at the baseline visit.

- 101 3. Study patient has any other active skin disease or condition (e.g., bacterial cellulitis, candida
102 intertrigo, extensive condyloma) that may, in the opinion of the investigator, interfere with the
103 assessment of HS.
- 104 4. Study patient has a diagnosis of sarcoidosis, systemic lupus erythematosus, or active IBD.
105 Note: study patients with a diagnosis of Crohn's disease or ulcerative colitis are allowed if
106 they have no active symptomatic disease at screening or baseline.
- 107 5. Study patient has a primary immunosuppressive condition, including taking
108 immunosuppressive therapy following an organ transplant, or has had a splenectomy.
- 109 Other medical conditions
- 110 6. Female study patient who is breastfeeding, pregnant, or plans to become pregnant during the
111 study or within 20 weeks following the final dose of IMP.
- 112 7. Study patient has an active infection or history of infection(s) as follows:
- 113 a. Any infection requiring systemic treatment within 14 days prior to baseline
- 114 b. A serious infection, defined as requiring hospitalisation or intravenous anti-
115 infective(s) within two months prior to the baseline visit
- 116 c. A history of opportunistic, recurrent, or chronic infections that, in the opinion of the
117 investigator, might cause this study to be detrimental to the study patient.
118 Opportunistic infections are infections caused by uncommon pathogens (e.g.,
119 *Pneumocystis jirovicii*, cryptococcosis), or unusually severe infections caused by
120 common pathogens (e.g., cytomegalovirus, herpes zoster)
- 121 8. Study patient has any of the following:
- 122 a. Known active tuberculosis (TB) disease
- 123 b. History of active TB involving any organ system unless adequately treated according
124 to World Health Organization/Centers for Disease Control and Prevention
125 therapeutic guidance and proven to be fully recovered upon consult with a TB
126 specialist
- 127 c. Latent TB infection (LTBI). Patients with LTBI diagnosed during screening must
128 have completed a course of prophylaxis prior to IMP dosing. Patients can be
129 rescreened after completion of a full course of prophylaxis plus a wash-out of least
130 five half-lives of the prophylactic medication(s) prior to baseline to avoid any
131 interference with the study efficacy measurements (e.g., concomitant antibiotics).
132 Prophylaxis should be in accordance with applicable clinical guidelines and TB
133 specialist judgment based on the origin of infection
- 134 d. High risk of exposure to TB infection
- 135 e. Current pulmonary nontuberculous mycobacterial (NTM) infection or history of
136 pulmonary NTM infection unless proven to be fully recovered
- 137 Note: For further information relating to definitions of known active TB, past history of TB, LTBI,
138 high risk of acquiring TB infection and NTM infection refer to Section 8.2.6 of the Study protocol.
- 139 9. Study patient has an acute or chronic hepatitis B virus, hepatitis C virus (HCV), or human
140 immunodeficiency virus (HIV) infection. Study patients who have evidence of, or tested
141 positive for, hepatitis B or hepatitis C will be excluded. A positive test for hepatitis B virus is
142 defined as: 1) positive for hepatitis B surface antigen, or 2) positive for anti-hepatitis B core
143 antibody. A positive test for HCV is defined as: 1) positive for hepatitis C antibody, and 2)
144 positive via a confirmatory test for HCV (for example, HCV polymerase chain reaction).
- 145 10. Study patients with concurrent malignancy are excluded. Study patients with a history of
146 malignancy within the past 5 years prior to the screening visit are excluded, EXCEPT if the
147 malignancy was a cutaneous squamous or basal cell carcinoma, or in situ cervical cancer that
148 has been treated and is considered cured.
- 149 11. Study patient has a history of a lymphoproliferative disorder including lymphoma or current
150 signs and symptoms suggestive of lymphoproliferative disease.

- 151 12. Study patient has had major surgery within the 3 months prior to the baseline visit, or has
152 planned major surgery after entering the study.
- 153 13. Study patient has any systemic disease (i.e., cardiovascular, neurological, renal, liver,
154 metabolic, gastrointestinal, haematological, immunological, etc.) considered by the
155 investigator to be uncontrolled, unstable, or likely to progress to a clinically significant degree
156 during the course of the study.
- 157 14. Study patient has had a myocardial infarction or stroke within the 6 months prior to the
158 screening visit.
- 159 15. Study patient has a history of chronic alcohol or drug abuse within 6 months prior to screening
160 as evaluated by the investigator based on medical history, interview, and/or results of the
161 screening urine drug screen.
- 162 16. Study patient has the presence of active suicidal ideation, or positive suicide behaviour using
163 the “Screening” version of the electronic Columbia Suicide Severity Rating Scale (eC-SSRS)
164 with either of the following criteria:
- 165 a. Study patient has a history of a suicide attempt within the 5 years prior to the
166 screening visit. Study patients with a history of a suicide attempt more than 5 years
167 ago should be evaluated by a mental healthcare practitioner before enrolling into the
168 study.
- 169 b. Suicidal ideation in the past month prior to the screening visit as indicated by a
170 positive response (“Yes”) to either Question 4 or Question 5 of the “Screening”
171 version of the eC-SSRS.
- 172 17. Study patient has presence of moderately severe major depression or severe major depression
173 indicated by a score of ≥ 15 using the screening Patient Health Questionnaire Depression
174 Module (PHQ-9). Medication used to treat depression should be stable for 8 weeks prior to
175 baseline.
- 176 18. Study patient has a known hypersensitivity to any components of bimekizumab or
177 comparative drugs as stated in the Study Protocol.
- 178 Previous/concomitant therapy
- 179 19. Study patient has had prior treatment with an IL-17 biologic response modifier or has
180 participated in IL-17 biologic response modifier study unless an appropriate washout has been
181 performed since the last dose of IMP (within 6 months prior to the baseline visit or five half-
182 lives [whichever is greater]).
- 183 20. Study patient received prescription topical therapies for the treatment of HS within 14 days
184 prior to the baseline visit.
- 185 21. Study patient is currently receiving systemic nonbiologic or biologic therapies for HS with
186 potential therapeutic impact for HS. Note: If study patient received systemic nonbiologic or
187 biologic therapies for HS and stopped these treatments, washout periods should be applied as
188 shown in Table 6–3 of the Study Protocol. Note: this does not apply to study patients who
189 may be eligible for randomisation into the antibiotic strata.
- 190 22. If study patient is using concomitant, non-opioid analgesics for HS-related or non-HS-related
191 pain as permitted by protocol, they should be on a stable (scheduled) dose for at least 14 days
192 prior to the baseline visit and anticipate continuing that dose through week 16 unless a
193 decrease in dose is warranted based on symptoms. Opioid analgesics are excluded. Note: As
194 needed (pro re nata [PRN]) use is not considered a stable dose, but (for example) taking a
195 nonsteroidal anti-inflammatory drug (NSAID) 3 times per week, every week is considered a
196 stable dose.
- 197 23. Study patient has received any live (including attenuated) vaccination within the 8 weeks prior
198 to the baseline visit (e.g., inactivated influenza and pneumococcal vaccines are allowed, but
199 nasal influenza vaccination is not permitted). Live vaccines are not allowed during the study,
200 including the safety follow-up (SFU) Period (20 weeks after the last dose of IMP).
- 201 24. Study patient has received Bacillus Calmette-Guerin vaccination within 1 year prior to IMP
202 administration.

Previous/concurrent clinical study experience

25. Study patient has previously participated in this study or study patient has previously been assigned to treatment in a study of the medication under investigation in this study, and received at least 1 dose of IMP (including placebo).
26. Study patient is currently participating in another study of a systemic medication under investigation, including SFU. Study patient must be washed out of the medication as indicated in Table 6–3 of the Study Protocol.
27. Study patient is currently participating in another study of a topical medication under investigation, including SFU. Study patient must be washed out of the medication for 4 weeks prior to the baseline visit.
28. Study patient is currently, or was within the 4 weeks prior to the baseline visit, participating in another study of a medical device under investigation.

Diagnostic assessments

29. Study patient has laboratory abnormalities at screening, including any of the following:
 - a. $\geq 3\times$ the upper limit of normal (ULN) of any of the following: alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP)
 - b. Bilirubin $>1.5\times$ ULN (isolated bilirubin $>1.5\times$ ULN is acceptable if bilirubin is fractionated and direct bilirubin $<35\%$)
 - c. White blood cell count $<3.00\times 10^3/\mu\text{L}$
 - d. Absolute neutrophil count $<1.5\times 10^3/\mu\text{L}$
 - e. Lymphocyte count <500 cells/ μL
 - f. Haemoglobin $<8.5\text{g/dL}$

Note: Individual screening tests for which the results are in error, borderline, or indeterminate for inclusion in the study can be repeated once for confirmation during the screening period. Upon retesting, study patients whose results remain outside this threshold should not be randomised.

30. Study patient has any other laboratory abnormality, which, in the opinion of the investigator, will prevent the study patient from completing the study or will interfere with the interpretation of the study results.

Other exclusions

31. Study patient is a UCB employee or is an employee of third-party organisations involved in the study.
32. Study patient and/or his or her immediate family member is an employee, volunteer, or other worker at the investigative site either affiliated or not affiliated with this study. Immediate family is defined as a spouse, parent, child, or sibling whether biological or legally adopted.

Independent DMC and subject matter expert adjudication committees

An independent data monitoring committee (DMC) periodically reviewed unblinded safety data on an ongoing basis to assess the risk of bimekizumab in study patients with moderate to severe HS. DMC membership included experienced clinicians and a statistician, all of whom had expertise in clinical studies. Separate Cardiovascular, Gastrointestinal, and Neuropsychiatric Adjudication Committees also periodically reviewed data from these studies.

All data provided are specified per the DMC charter. Unblinded efficacy data analysis was provided to the DMC when approximately 40% of study patients in BE HEARD I and approximately 66% of study patients in BE HEARD II had reached the primary efficacy endpoint at Week 16. At this analysis, formal primary endpoint comparisons between active treatment groups and placebo were provided so that the DMC could make an assessment regarding the futility of each dose group. A priori, no early stopping of a treatment group or study on the basis of superior efficacy was planned. Therefore, no type I error adjustment was required.

Both DMC and Adjudication Committee members did not participate in the study as principal or co-investigators, or as study patient care physicians, and were also not members of the study team at UCB or the conducting clinical research organisation. The duration of membership for the committees was inclusive of planned analyses for this study.

Measures to ensure that unblinded interim results are not disseminated beyond the DMC will be implemented. Details of futility stopping boundaries and decision rules will be detailed in the DMC Charter.

Clinical definitions

The lesion count was defined as an assessment of all the various skin “appearances” termed “lesions” in HS study patients. The lesion count included the following:

- Abscesses (circumscribed collection of purulent exudate frequently associated with swelling and other signs of inflammation, such as fluctuance, tenderness, and pain)
- Draining tunnels (fistulas/sinus tracts) (pathologic passageway connecting to the skin surface from dermis or subcutaneous tissue/pathologic passageway that develops into a channel to the skin surface that drains serous or purulent fluid, either spontaneously or by gentle palpation)
- Non draining tunnels (fistulas/sinus tracts) (pathologic passageway connecting to the skin surface from dermis or subcutaneous tissue/pathologic passageway that develops into a channel to the skin surface that does not drain serous or purulent fluid)
- Noninflammatory nodules (nontender or minimally tender, non-erythematous nodules)
- Inflammatory nodules (a tender, erythematous, well defined nodule. The lesion has no evidence of fluctuance. A pyogenic granuloma lesion is considered an inflammatory nodule; a papule or pustule is not considered an inflammatory nodule)
- Scars of HS lesions (enlargement or overgrowth of a scar so that it extends above the surrounding skin surface)

Statistical analyses

A sensitivity analysis of the primary endpoint was performed based on the full analysis set, which consisted of all study participants who received at least one full or partial dose of study treatment and had a valid baseline (week 0) and post-baseline measurement (any study visit after baseline [week 0]) for AN, DT, and inflammatory nodule counts.

The logistic regression model used for the primary and secondary endpoint analyses used fixed effects of treatment, baseline Hurley Stage, and baseline antibiotic use, and analgesic use for HSSDD worst skin pain response.

The analysis of covariance model used for analysis of continuous secondary endpoints used fixed effects of treatment, baseline Hurley Stage, baseline antibiotic use, and baseline DLQI or HSSDD worst skin pain score as factors; additionally analgesic use for HSSDD worst skin pain response.

The multiple imputation procedure was as follows: intermittent missing data were imputed using the Markov Chain Monte Carlo (MCMC) method. Monotone missing data were imputed using monotone regression where Hurley Stage at baseline, baseline antibiotic use, and value of the variable of interest at baseline and at each post-baseline visit (prior to the time point of interest) were included as predictors in the imputation model. The post-baseline values were specified in chronological order in the imputation model so that SAS® PROC MI imputed variables from left to right (e.g., the week 2 value was first imputed using regression based on the baseline value, and then the week 4 value was imputed using regression based on baseline and week 2 values, etc.). The analysis was then based on 100 imputations. The imputation procedure was performed for each treatment regimen. A data set was created for each treatment group of participants with observed values and those needing estimation by multiple imputation.

301 Exploratory endpoints in BE HEARD I and BE HEARD II

Objectives	Endpoints
Evaluate the efficacy of bimekizumab on HiSCR, other HS Scores, and other clinical measures of disease activity at various timepoints in study patients with moderate to severe HS	<ul style="list-style-type: none"> • Time to response of HiSCR25, HiSCR50, HiSCR75, HiSCR90, and HiSCR100 • HiSCR25, HiSCR50, HiSCR75, HiSCR90, and HiSCR100 • Absolute change from Baseline in International Hidradenitis Suppurativa Severity Score System • Change from Baseline in the HS-Physician's Global Assessment 6-point scale • Absolute and percentage change from baseline in hs-CRP • Initiation of systemic antibiotic rescue therapy • HiSCR25, HiSCR50, HiSCR75, HiSCR90, and HiSCR100 at both weeks 16 and 48 • Time to loss of response of HiSCR50, HiSCR75, HiSCR90, and HiSCR100 in week 16 responders • Partial responders (defined as a $\geq 25\%$ reduction in AN count from baseline) at week 16 who progress to HiSCR50 during the maintenance treatment period
Evaluate the efficacy of bimekizumab on abscesses, nodules, and draining tunnels at various timepoints in study patients with moderate to severe HS	<ul style="list-style-type: none"> • Change and percentage change from Baseline in lesion counts (abscess count, inflammatory nodule count, AN count, and draining tunnel count) • AN count of 0, 1, or 2 • AN25, AN50, AN75, AN90, AN100 • Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to baseline) by week 16 (BE HEARD I only) • Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48 • Time to flare from weeks 0 to 16 • Time to flare from week 16 to 48
Evaluate the efficacy of bimekizumab on patient-reported outcome measures at various timepoints in study patients with moderate to severe HS	<ul style="list-style-type: none"> • Absolute and percentage change (worst and average pain) from baseline in HS Skin Pain score (11-point numeric rating scale) • Pain response (defined as a decrease from baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) • Pain response (defined as a decrease from baseline in HSSQ weekly worst skin pain score at or beyond the threshold for clinically meaningful change) • Pain response (at least a 30% reduction and at least a 1-unit reduction from baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study patients with a score of ≥ 3 at baseline • Pain response (at least a 30% reduction and at least a 1-unit reduction from baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study patients with a score of ≥ 3 at baseline • Pain response (at least a 30% reduction and at least a 1-unit reduction from baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study patients with a score of ≥ 3 at baseline • Absolute change from baseline in DLQI Total Score DLQI Total Score of 0 or 1 • Minimum clinically important difference (improvement from baseline of 4 or more) in the DLQI Total Score among study patients with a baseline score of at least 4) • Absolute change from baseline in HiSQOL[®] domain scores (symptoms, psychosocial, activities, adaptations) • Absolute change from baseline in Patient Global Impression of HS Severity • Absolute change from Baseline in Patient Global Impression of Severity of HS Skin Pain • Absolute change from Baseline in each of the other HS symptoms – itch, drainage or oozing of HS lesions, and smell or odour • Response on other HS symptoms (11-point numeric rating scale) – itch, drainage or oozing of HS lesions, and smell or odour • Responses to the EQ-5D-3L, absolute and changes from baseline in EQ-5D-3L visual analogue scale scores • Absolute change from baseline in Work Productivity and Activity Impairment Questionnaire-Specific Health Problem v2.0 adapted to HS scores • Domain Scores (effectiveness, convenience and global satisfaction) on the Treatment Satisfaction Questionnaire – Medication-9
Evaluate the effect of bimekizumab on other safety measures at various	<ul style="list-style-type: none"> • Adverse events of special interest (Hy's Law) • Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation

timepoints in study patients with moderate to severe HS	<p>and behaviour, major adverse cardiovascular events, hepatic events and potential drug-induced liver injury (PDILI), malignancies, and inflammatory bowel disease</p> <ul style="list-style-type: none"> • Absolute change from baseline in the PHQ-9 score • Absolute change from baseline in vital signs • Absolute change from baseline in clinical laboratory values (chemistry and haematology) • ECG results
---	--

302

303

304 **Supplementary Results**

305 **Primary efficacy outcome in BE HEARD I and BE HEARD II at week 16 using OC data**

306 In BE HEARD I, 151 of 257 (59%) patients in the bimekizumab every 2 weeks group versus 24 of 65
307 (37%) patients in the placebo group achieved HiSCR50 (OR 2.32 [97.5% CI 1.13–4.77]; $p=0.009$). In
308 BE HEARD II, 164 of 265 (62%) patients in the bimekizumab every 2 weeks group versus 24 of 70
309 (34%) patients in the placebo group achieved HiSCR50 (OR 3.24 [1.66–6.30]; $p<0.001$).

310

311 In BE HEARD I, 72 of 124 (58%) patients in the bimekizumab every 4 weeks group versus 24 of 65
312 (37%) patients in the placebo group achieved HiSCR50 (OR 2.16 [97.5% CI 0.98–4.75]; $p=0.029$; not
313 met). In BE HEARD II, 80 of 133 (60%) patients in the bimekizumab every 4 weeks group versus 24
314 of 70 (34%) patients in the placebo group achieved HiSCR50 (OR 2.96 [1.43–6.14]; $p<0.001$).

315 **Supplementary Tables**

316 **Table S1. Number of patients with HiSCR50 data observed or imputed at week 16 in BE HEARD I and BE HEARD II**

Week 16	BE HEARD I			BE HEARD II		
	Placebo (N=72)	Bimekizumab 320 mg Q2W (N=289)	Bimekizumab 320 mg Q4W (N=144)	Placebo (N=74)	Bimekizumab 320 mg Q2W (N=291)	Bimekizumab 320 mg Q4W (N=144)
OC	51 (70.8%)	214 (74.0%)	102 (70.8%)	64 (86.5%)	225 (77.3%)	116 (80.6%)
NRI – Antibiotic use	15 (20.8%)	47 (16.3%)	22 (15.3%)	6 (8.1%)	45 (15.5%)	18 (12.5%)
NRI – Discontinued due to adverse event	1 (1.4%)	5 (1.7%)	6 (4.2%)	1 (1.4%)	8 (2.7%)	1 (0.7%)
NRI – Discontinued due to loss of efficacy	0	0	0	0	0	0
MI	5 (6.9%)	23 (8.0%)	14 (9.7%)	3 (4.1%)	13 (4.5%)	9 (6.3%)

317 Randomised set. HiSCR: Hidradenitis Suppurativa Clinical Response; HiSCR50: $\geq 50\%$ reduction from baseline in the total abscess and inflammatory nodule count with no
318 increase from baseline in abscess or draining tunnel count; MI: multiple imputation; mNRI: modified non-responder imputation; NRI: non-responder imputation;
319 OC: observed case; Q2W: every 2 weeks; Q4W: every 4 weeks.

320 **Table S2. HiSCR90 and HiSCR100 responses at week 16 and week 48**

		BE HEARD I				BE HEARD II			
Endpoint (mNRI: % [95% CI]; OC: % [n/Nsub])		Bimekizumab 320 mg Q2W/Q2W (N=143)	Bimekizumab 320 mg Q2W/Q4W (N=146)	Bimekizumab 320 mg Q4W/Q4W (N=144)	Placebo/ Bimekizumab 320 mg Q2W (N=72)	Bimekizumab 320 mg Q2W/Q2W (N=145)	Bimekizumab 320 mg Q2W/Q4W (N=146)	Bimekizumab 320 mg Q4W/Q4W (N=144)	Placebo/ Bimekizumab 320 mg Q2W (N=74)
HiSCR90									
Week 16	mNRI (HS-ABX)	21.9 (14.8, 29.0)	20.6 (13.8, 27.4)	20.1 (13.3, 26.9)	11.3 (3.9, 18.6)	19.9 (13.3, 26.6)	23.2 (16.2, 30.2)	21.0 (14.1, 27.8)	5.8 (0.3, 11.3)
	OC	22.2 (28/126)	21.4 (28/131)	20.2 (25/124)	12.3 (8/65)	21.1 (28/133)	24.2 (32/132)	22.6 (30/133)	7.1 (5/70)
Week 48	mNRI (HS-ABX)	29.2 (21.0, 37.4)	29.6 (21.6, 37.6)	25.3 (17.2, 33.3)	26.1 (14.8, 37.3)	31.8 (23.6, 39.9)	35.9 (27.7, 44.1)	36.7 (28.3, 45.0)	39.8 (28.1, 51.5)
	OC	41.2 (40/97)	36.5 (38/104)	33.7 (29/86)	34.1 (15/44)	38.2 (42/110)	46.7 (50/107)	48.6 (53/109)	44.3 (27/61)
HiSCR100									
Week 16	mNRI (HS-ABX)	14.6 (11.1, 18.2)	17.2 (13.6, 20.8)	15.0 (10.6, 19.4)	7.0 (1.1, 12.9)	16.6 (13.0, 20.2)	16.1 (12.6, 19.6)	16.5 (12.1, 21.0)	4.3 (0.0, 9.0)
	OC	14.3 (18/126)	17.6 (23/131)	14.5 (18/124)	7.7 (5/65)	16.5 (22/133)	16.7 (22/132)	18.0 (24/133)	4.3 (3/70)
Week 48	mNRI (HS-ABX)	22.4 (17.0, 27.8)	22.9 (17.7, 28.0)	20.2 (14.9, 25.4)	17.8 (8.1, 27.5)	24.9 (19.6, 30.2)	24.3 (19.1, 29.5)	26.9 (21.5, 32.3)	34.1 (22.8, 45.3)
	OC	30.9 (30/97)	26.9 (28/104)	26.7 (23/86)	22.7 (10/44)	29.1 (32/110)	30.8 (33/107)	33.9 (37/109)	37.7 (23/61)

321 Randomised set. Data were imputed using mNRI (HS-ABX): patients who took systemic antibiotics identified as rescue medication for HS by the principal investigator or
322 who discontinued due to adverse events or lack of efficacy were treated as non-responders at all subsequent visits. Other missing data were imputed via MI. OC: all available
323 data after an intercurrent event were summarised as recorded in the database, and all missing data were left missing. ABX: antibiotics; AN: abscess and inflammatory nodule;
324 CI: confidence interval; DT: draining tunnel; HiSCR: Hidradenitis Suppurativa Clinical Response; HiSCR90: ≥90% reduction from baseline in the total abscess and
325 inflammatory nodule count with no increase from baseline in abscess or DT count; HiSCR100: 100% reduction from baseline in the total AN count with no increase from
326 baseline in abscess or DT count; HS: hidradenitis suppurativa; mNRI: modified non-responder imputation; OC: observed case; Q2W: every 2 weeks; Q4W: every 4 weeks.
327

328 **Table S3. AN and DT count at baseline, week 16 and week 48**

Absolute value (Mean ± SE)	BE HEARD I				BE HEARD II			
	Bimekizumab 320 mg Q2W/Q2W (N=143)	Bimekizumab 320 mg Q2W/Q4W (N=146)	Bimekizumab 320 mg Q4W/Q4W (N=144)	Placebo/ Bimekizumab 320 mg Q2W (N=72)	Bimekizumab 320 mg Q2W/Q2W (N=145)	Bimekizumab 320 mg Q2W/Q4W (N=146)	Bimekizumab 320 mg Q4W/Q4W (N=144)	Placebo/ Bimekizumab 320 mg Q2W (N=74)
AN count								
Baseline	13.7 ± 0.8	16.9 ± 1.3	17.8 ± 2.1	15.0 ± 1.4	15.8 ± 1.1	17.5 ± 1.4	17.6 ± 1.3	13.9 ± 0.9
Week 16	6.0 ± 0.6	9.1 ± 1.1	8.7 ± 1.0	12.0 ± 1.7	7.8 ± 1.0	7.5 ± 1.1	9.0 ± 1.2	9.5 ± 1.0
Week 48	4.1 ± 0.5	4.8 ± 0.5	6.1 ± 0.8	3.9 ± 0.7	4.4 ± 0.6	4.1 ± 0.7	5.9 ± 1.1	2.9 ± 0.5
DT count								
Baseline	4.0 ± 0.4	4.0 ± 0.4	3.8 ± 0.4	3.2 ± 0.5	3.6 ± 0.3	3.6 ± 0.3	2.8 ± 0.3	3.5 ± 0.4
Week 16	2.0 ± 0.3	2.7 ± 0.4	2.2 ± 0.3	2.8 ± 0.5	2.1 ± 0.3	1.7 ± 0.3	1.6 ± 0.2	3.1 ± 0.4
Week 48	1.5 ± 0.2	2.0 ± 0.3	1.5 ± 0.2	1.3 ± 0.3	1.7 ± 0.3	1.5 ± 0.2	1.2 ± 0.2	1.6 ± 0.3

329 Randomised set. Data were imputed using MI (HS-ABX): patients who took systemic antibiotics identified as rescue medication for HS by the principal investigator or who
330 discontinued study treatment due to an adverse event or lack of efficacy were treated as missing and then subsequently imputed using MI. All other missing data were also
331 imputed using MI. ABX: antibiotics; AN: abscess and inflammatory nodule; DT: draining tunnel; HS: hidradenitis suppurativa; MI: multiple imputation; Q2W: every 2
332 weeks; Q4W: every 4 weeks; SE: standard error.

Table S4. TEAEs with BKZ treatment to week 16 in BE HEARD I and BE HEARD II

Week 16	BE HEARD I	BE HEARD II
n (%)	Bimekizumab Total (N=429) 100 PY=1·30	Bimekizumab Total (N=432) 100 PY=1·32
Any TEAE	286 (66·7%)	260 (60·2%)
Serious TEAE	10 (2·3%)	12 (2·8%)
Discontinuation due to TEAE	16 (3·7%)	15 (3·5%)
Drug-related TEAE	121 (28·2%)	143 (33·1%)
Severe TEAE	11 (2·6%)	17 (3·9%)
Deaths	0	0
Top three most common TEAEs^a		
Hidradenitis	31 (7·2%)	38 (8·8%)
Oral candidiasis	19 (4·4%)	29 (6·7%)
Diarrhoea	30 (7·0%)	23 (5·3%)
Headache	30 (7·0%)	25 (5·8%)
TEAEs of interest		
Infections and infestations	150 (35·0%)	134 (31·0%)
Serious infections	1 (0·2%)	0
Opportunistic infections ^b	2 (0·5%)	1 (0·2%)
Fungal infections	51 (11·9%)	59 (13·7%)
<i>Candida</i> infections	29 (6·8%)	41 (9·5%)
Oral candidiasis	19 (4·4%)	29 (6·7%)
Neutropenia	0	0
Hypersensitivity reaction (SMQ, narrow) ^c	42 (9·8%)	41 (9·5%)
Dermatitis and eczema	20 (4·7%)	29 (6·7%)
Serious hypersensitivity reaction	0	0
Adjudicated suicidal ideation and behaviour	2 (0·5%)	1 (0·2%)
Adjudicated major adverse cardiovascular events	0	0
Hepatic events	10 (2·3%)	9 (2·1%)
>5x ULN elevation of AST/ALT ^d	3 (0·7%) ^e	0 ^f
Malignancies	0	1 (0·2%)
Definite or probable adjudicated inflammatory bowel disease	1 (0·2%)	3 (0·7%)

Safety set, MedDRA (version 19·0). Hepatic events category includes events in the SMQ "Drug related hepatic disorders - comprehensive search (SMQ)", excluding the following two sub-SMQs: "Liver neoplasms, benign (incl. cysts and polyps) (SMQ)" and "Liver neoplasms, malignant and unspecified (SMQ)"; hepatic events category includes all post-baseline assessments including those at unscheduled visits but excluding any that occur more than 140 days after the last administration of study medication, counting a patient only once. ^aValues in bold are the top three most common TEAEs for bimekizumab treatment groups in each study (BE HEARD I: hidradenitis, headache, diarrhoea; BE HEARD II: hidradenitis, oral candidiasis, headache); ^bOpportunistic infections were localised mucocutaneous events, as defined by internal company conventions; ^cThere were no incidences of anaphylactic reactions; ^dNo elevations >5x ULN were adjudicated to be highly likely or definitely related to bimekizumab; ^en=424; ^fn=430. ALT: alanine aminotransferase; AST: aspartate aminotransferase; MedDRA: Medical Dictionary for Regulatory Activities; PY: patient-year;

346 TEAE: treatment-emergent adverse event; SMQ: Standardised MedDRA Query; ULN: upper limit of
347 normal.

348 **Table S5. TEAEs by severity grade category to week 16 in BE HEARD I and BE HEARD II**

Week 16	BE HEARD I		BE HEARD II	
n (%)	Bimekizumab Total (N=429) 100 PY=1·30	Placebo (N=72) 100 PY=0·22	Bimekizumab Total (N=432) 100 PY=1·32	Placebo (N=74) 100 PY=0·23
Mild	182 (42·4%)	28 (38·9%)	143 (33·1%)	26 (35·1%)
Moderate	93 (21·7%)	20 (27·8%)	100 (23·1%)	14 (18·9%)
Severe	11 (2·6%)	0	17 (3·9%)	2 (2·7%)

349 Safety set, MedDRA (version 19·0). MedDRA: Medical Dictionary for Regulatory Activities;
350 PY: patient-year; TEAE: treatment-emergent adverse event.

351 **Table S6. Severe grade TEAEs to week 16 in BE HEARD I and II**

BE HEARD I	BE HEARD II
Any bimekizumab treatment regimen	Any bimekizumab treatment regimen
11 patients reported severe TEAEs: ^a <ul style="list-style-type: none"> • Gastrointestinal disorders: 1 • General disorders and administration site conditions: 1 • Hepatobiliary disorders: 1 • Infections and infestations: 2 • Investigations: 2 • Pregnancy, puerperium and perinatal conditions: 1 • Psychiatric disorders: 1 • Skin and subcutaneous tissue disorders: 2 • Vascular disorders: 1 	17 patients reported severe TEAEs: ^b <ul style="list-style-type: none"> • Blood and lymphatic system disorders: 1 • Gastrointestinal disorders: 1 • Hepatobiliary disorders: 1 • Infections and infestations: 3 • Musculoskeletal and connective tissue disorders: 2 • Neoplasms benign, malignant and unspecified (incl cysts and polyps): 1 • Nervous system disorders: 1 • Reproductive system and breast disorders: 1 • Skin and subcutaneous tissue disorders: 7
Placebo treatment	Placebo treatment
0	2 patients reported severe TEAEs: <ul style="list-style-type: none"> • General disorders and administration site conditions: 1 • Psychiatric disorders: 1

352 TEAEs presented by MedDRA (version 19·0) System Organ Class categories. ^a14 events in 11
353 patients; ^b19 events in 17 patients. MedDRA: Medical Dictionary for Regulatory Activities;
354 TEAE: treatment-emergent adverse event.

Table S7. TEAEs by treatment regimen to week 48 in BE HEARD I and BE HEARD II

Week 48	BE HEARD I			BE HEARD II		
n (%)	Bimekizumab 320 mg Q2W/Q2W (N=141) 100 PY=1·21	Bimekizumab 320 mg Q2W/Q4W (N=145) 100 PY=1·22	Bimekizumab 320 mg Q4W/Q4W (N=143) 100 PY=1·17	Bimekizumab 320 mg Q2W/Q2W (N=144) 100 PY=1·22	Bimekizumab 320 mg Q2W/Q4W (N=146) 100 PY=1·24	Bimekizumab 320 mg Q4W/Q4W (N=142) 100 PY=1·24
Any TEAE	126 (89·4%)	124 (85·5%)	122 (85·3%)	122 (84·7%)	128 (87·7%)	113 (79·6%)
Serious TEAE	13 (9·2%)	8 (5·5%)	13 (9·1%)	10 (6·9%)	5 (3·4%)	7 (4·9%)
Discontinuation due to TEAE	8 (5·7%)	10 (6·9%)	13 (9·1%)	9 (6·3%)	10 (6·8%)	8 (5·6%)
Drug-related TEAE	66 (46·8%)	72 (49·7%)	65 (45·5%)	72 (50·0%)	77 (52·7%)	50 (35·2%)
Severe TEAE	17 (12·1%)	7 (4·8%)	11 (7·7%)	13 (9·0%)	8 (5·5%)	14 (9·9%)
Deaths	1 (0·7%)	0	0	0	0	0
Top three most common TEAEs^a						
Hidradenitis	22 (15·6%)	27 (18·6%)	36 (25·2%)	26 (18·1%)	31 (21·2%)	26 (18·3%)
Coronavirus infection	23 (16·3%)	23 (15·9%)	16 (11·2%)	11 (7·6%)	15 (10·3%)	6 (4·2%)
Oral candidiasis	15 (10·6%)	16 (11·0%)	13 (9·1%)	22 (15·3%)	25 (17·1%)	14 (9·9%)
Diarrhoea	15 (10·6%)	12 (8·3%)	16 (11·2%)	8 (5·6%)	16 (11·0%)	10 (7·0%)
Headache	12 (8·5%)	16 (11·0%)	10 (7·0%)	13 (9·0%)	14 (9·6%)	13 (9·2%)
TEAEs of interest						
Infections and infestations	91 (64·5%)	89 (61·4%)	87 (60·8%)	85 (59·0%)	85 (58·2%)	76 (53·5%)
Serious infections	5 (3·5%)	3 (2·1%)	2 (1·4%)	2 (1·4%)	1 (0·7%)	1 (0·7%)
Opportunistic infections ^b	1 (0·7%)	3 (2·1%)	3 (2·1%)	1 (0·7%)	1 (0·7%)	2 (1·4%)
Fungal infections	33 (23·4%)	32 (22·1%)	35 (24·5%)	40 (27·8%)	39 (26·7%)	35 (24·6%)
<i>Candida</i> infections	20 (14·2%)	21 (14·5%)	22 (15·4%)	27 (18·8%)	29 (19·9%)	26 (18·3%)
Oral candidiasis	15 (10·6%)	16 (11·0%)	13 (9·1%)	22 (15·3%)	25 (17·1%)	14 (9·9%)

Week 48	BE HEARD I			BE HEARD II		
n (%)	Bimekizumab 320 mg Q2W/Q2W (N=141) 100 PY=1·21	Bimekizumab 320 mg Q2W/Q4W (N=145) 100 PY=1·22	Bimekizumab 320 mg Q4W/Q4W (N=143) 100 PY=1·17	Bimekizumab 320 mg Q2W/Q2W (N=144) 100 PY=1·22	Bimekizumab 320 mg Q2W/Q4W (N=146) 100 PY=1·24	Bimekizumab 320 mg Q4W/Q4W (N=142) 100 PY=1·24
Neutropenia	0	1 (0·7%)	0	0	0	0
Hypersensitivity reaction (SMQ, narrow) ^c	38 (27·0%)	30 (20·7%)	26 (18·2%)	23 (16·0%)	28 (19·2%)	23 (16·2%)
Dermatitis and eczema	22 (15·6%)	20 (13·8%)	15 (10·5%)	14 (9·7%)	20 (13·7%)	17 (12·0%)
Serious hypersensitivity reaction	0	0	0	1 (0·7%)	0	0
Adjudicated suicidal ideation and behaviour	1 (0·7%)	0	3 (2·1%)	0	1 (0·7%)	0
Adjudicated major adverse cardiovascular events	2 (1·4%)	1 (0·7%)	0	0	0	0
Hepatic events	12 (8·5%)	7 (4·8%)	2 (1·4%)	3 (2·1%)	7 (4·8%)	7 (4·9%)
>5x ULN elevation of AST/ALT ^d	2 (1·4%) ^e	2 (1·4%) ^f	0 ^e	1 (0·7%) ^g	0 ^h	2 (1·4%)
Malignancies	0	0	0	2 (1·4%)	0	1 (0·7%)
Definite or probable adjudicated inflammatory bowel disease	0	1 (0·7%)	1 (0·7%)	1 (0·7%)	1 (0·7%)	2 (1·4%)

Active medication set, MedDRA (version 19·0). Hepatic events category includes events in the SMQ "Drug related hepatic disorders - comprehensive search (SMQ)", excluding the following two sub-SMQs: "Liver neoplasms, benign (incl. cysts and polyps) (SMQ)" and "Liver neoplasms, malignant and unspecified (SMQ)"; hepatic events category includes all post-baseline assessments including those at unscheduled visits but excluding any that occur more than 140 days after the last administration of study medication, counting a patient only once. ^aValues in bold are the top three most common TEAEs for bimekizumab treatment groups in each study (BE HEARD I: hidradenitis, coronavirus infection, diarrhoea; BE HEARD II: hidradenitis, oral candidiasis, headache); ^bOpportunistic infections were localised mucocutaneous events, as defined by internal company conventions; ^cThere were no incidences of anaphylactic reactions related to bimekizumab; ^dNo elevations >5× ULN were adjudicated to be highly likely or definitely related to bimekizumab; ^en=140; ^fn=144; ^gn=143; ^hn=145. ALT: alanine aminotransferase; AST: aspartate aminotransferase; MedDRA: Medical Dictionary for Regulatory Activities; PY: patient-year; Q2W: every 2 weeks; Q4W: every 4 weeks; TEAE: treatment-emergent adverse event; SMQ: Standardised MedDRA Query; ULN: upper limit of normal.

Supplementary Figures

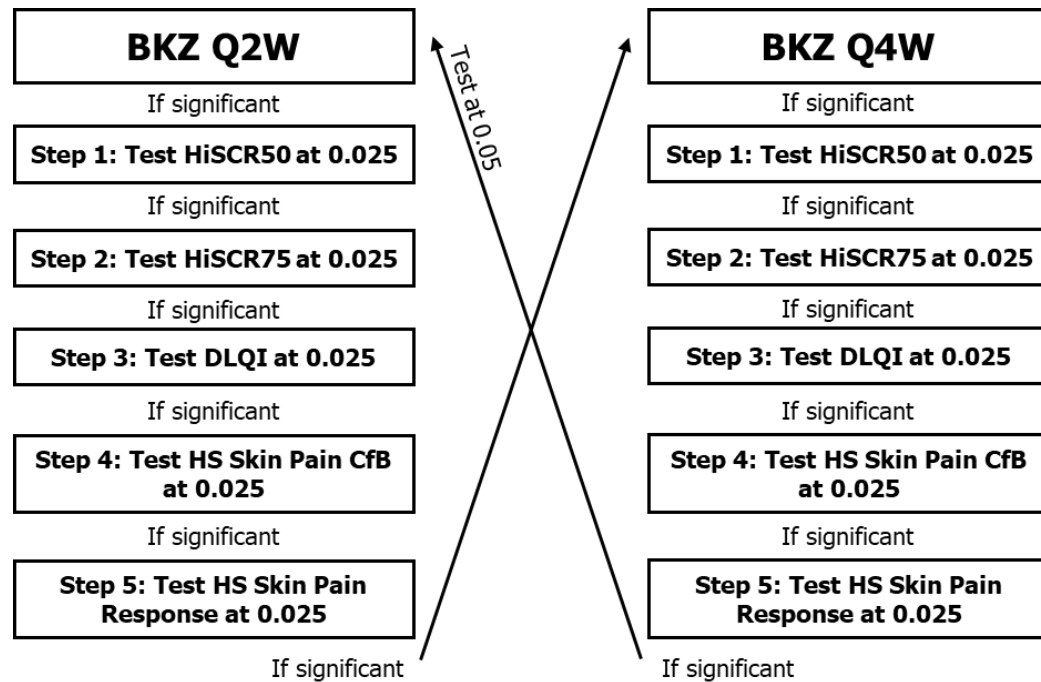
Figure S1. Study design



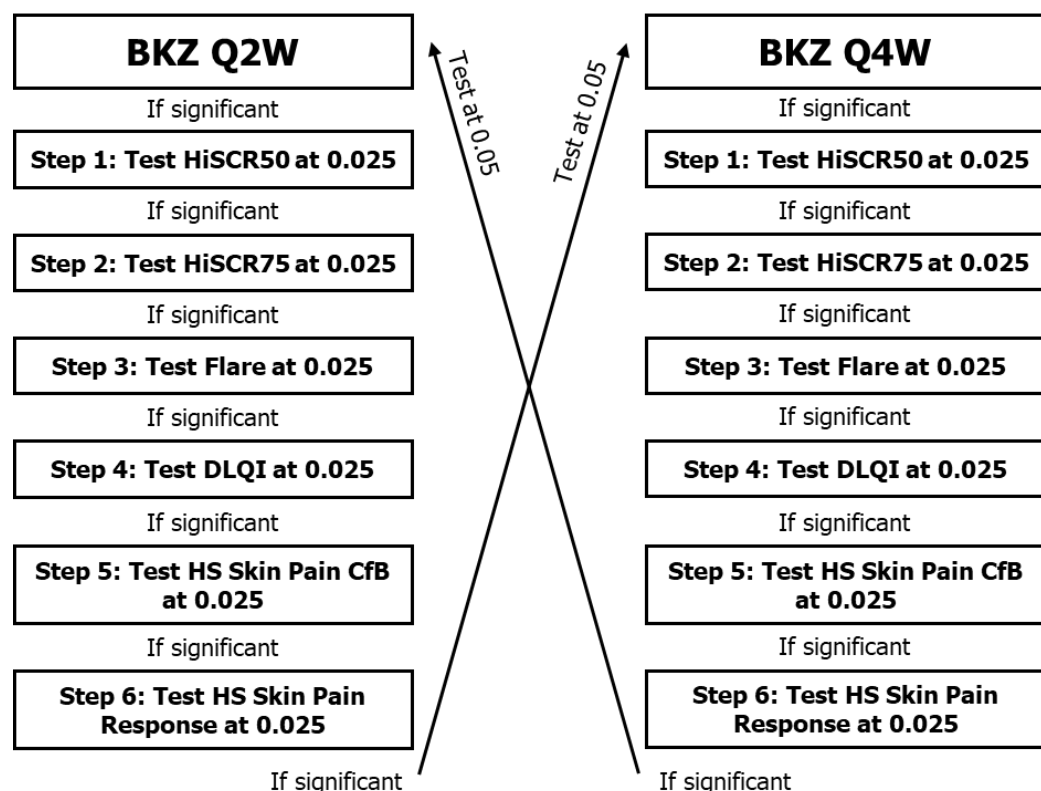
At baseline, patients with moderate to severe HS were randomised 2:2:2:1 to bimekizumab 320 mg every 2 weeks (Q2W) to week 48, bimekizumab 320 mg every 4 weeks (Q4W) to week 48, bimekizumab 320 mg Q2W to week 16 then bimekizumab 320 mg Q4W to week 48, or placebo to week 16 then bimekizumab 320 mg Q2W to week 48. AN: abscess and inflammatory nodule; BKZ: bimekizumab; DT: draining tunnel; HiSCR: Hidradenitis Suppurativa Clinical Response; HiSCR50: $\geq 50\%$ reduction from baseline in the total AN count with no increase from baseline in abscess or DT count; HS: hidradenitis suppurativa; Q2W: every 2 weeks; Q4W: every 4 weeks.

Figure S2. Sequence of testing

A) BE HEARD I



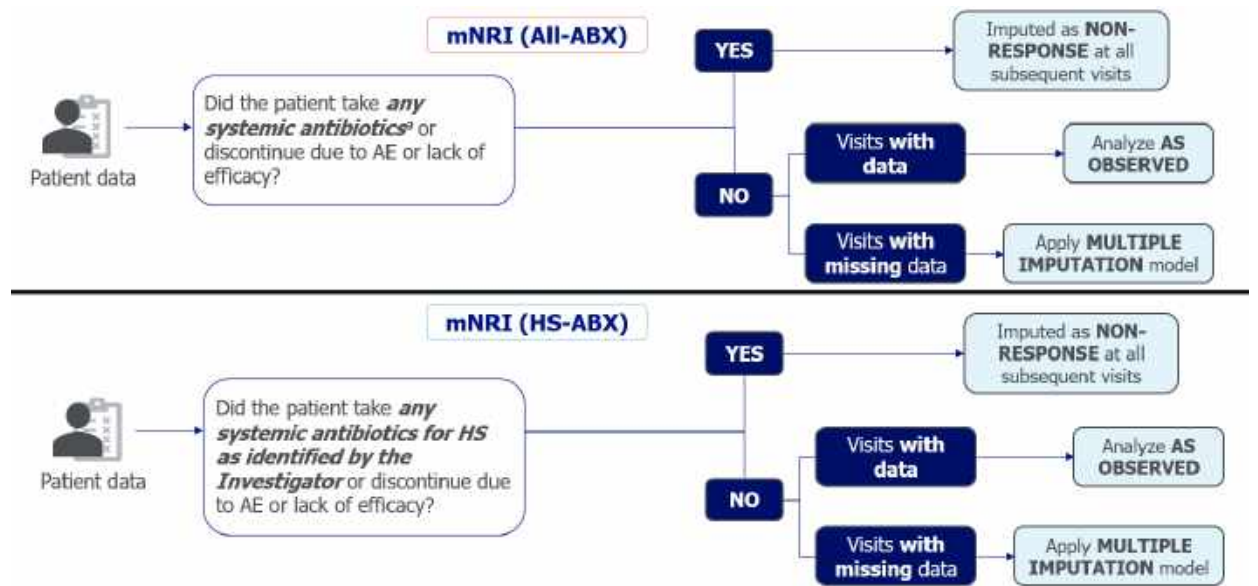
B) BE HEARD II



HS skin pain response was tested among study patients with a score of ≥ 3 at baseline. AN: abscess and inflammatory nodule; BKZ: bimekizumab; CFB: change from baseline; DLQI: Dermatology Life Quality Index; DT: draining tunnel; HiSCR: Hidradenitis Suppurativa Clinical Response;

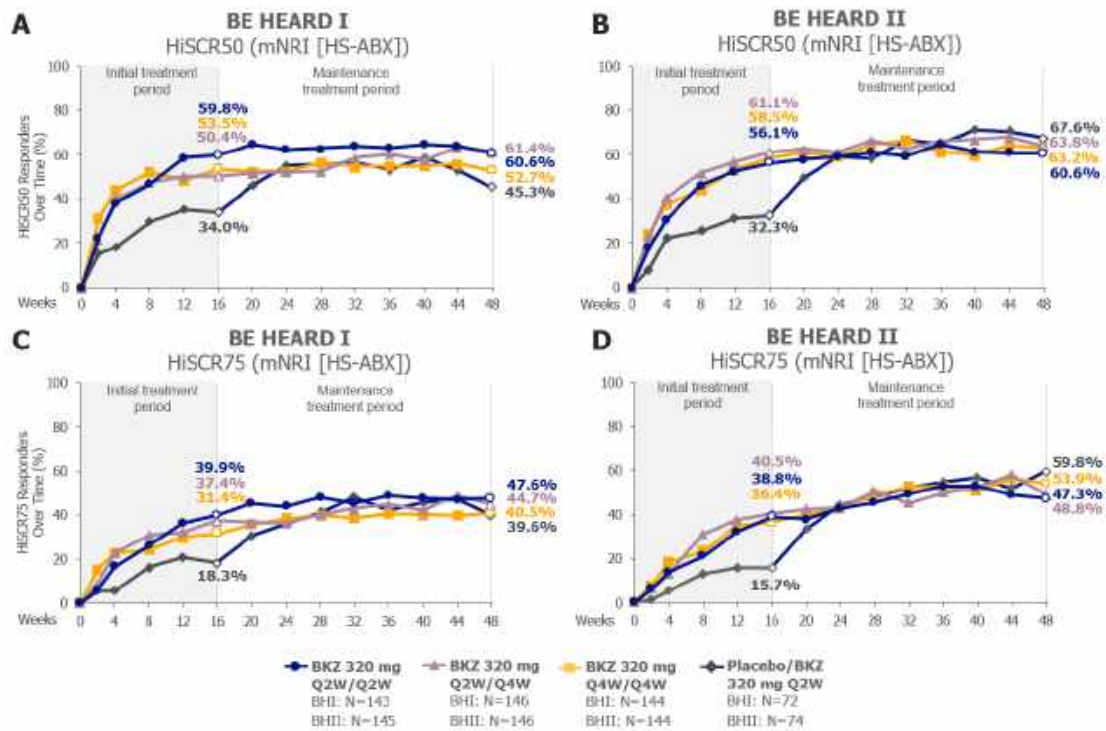
384 HiSCR50/75: $\geq 50/75\%$ reduction from baseline in the total AN count with no increase from baseline in
385 abscess or DT count; HS: hidradenitis suppurativa; Q2W: every 2 weeks; Q4W: every 4 weeks.

Figure S3. Missing data handling



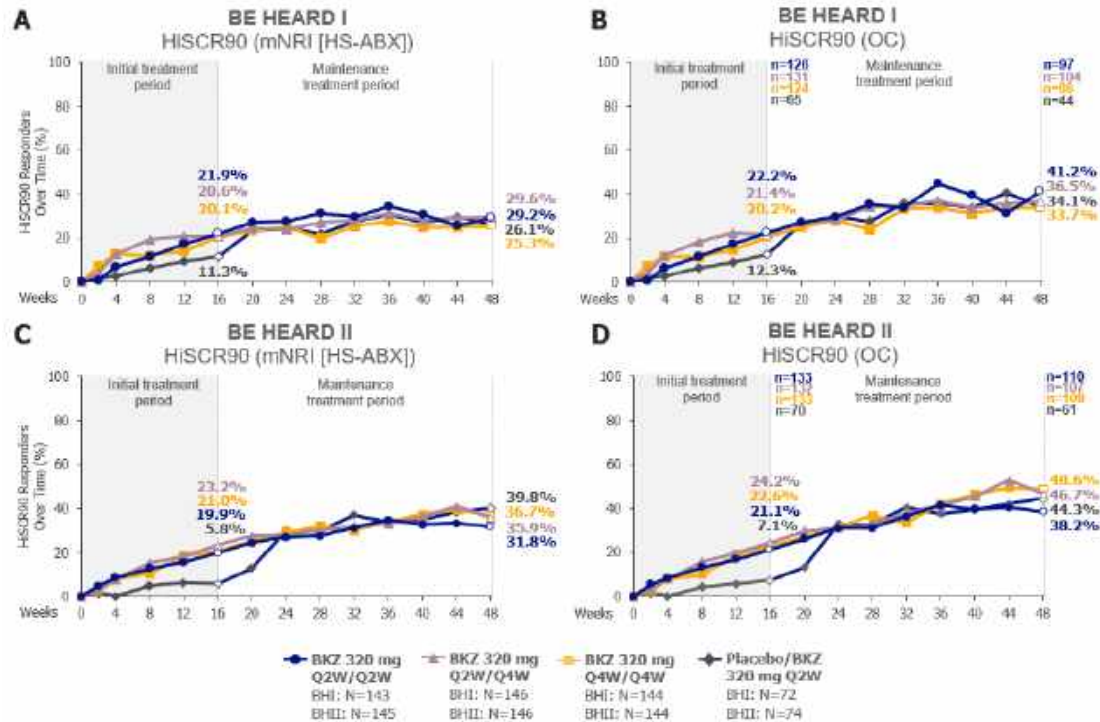
^aThis refers to any new systemic antibiotic for those who were not on systemic antibiotics at baseline and any change in dose for those who were on systemic antibiotics at baseline. ABX: antibiotics; AE: adverse event; HS: hidradenitis suppurativa; mNRI: modified non-responder imputation.

Figure S4. HiSCR50 and HiSCR75 responses over time to week 48



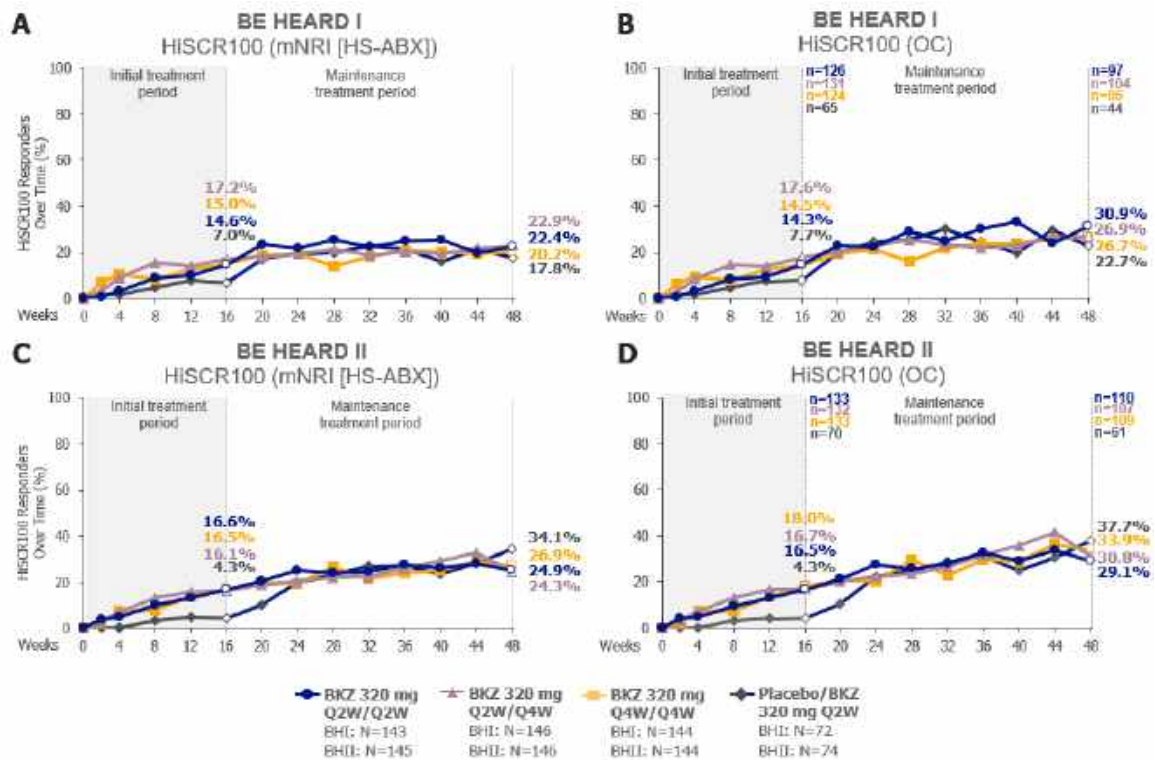
Achievement of HiSCR50 (A, B) and HiSCR75 (C, D) over time to week 48 in BE HEARD I and BE HEARD II. mNRI (HS-ABX): patients who took systemic antibiotics identified as rescue medication for HS by the principal investigator or who discontinued due to adverse events or lack of efficacy were treated as non-responders at all subsequent visits. Other missing data were imputed via MI. ABX=antibiotics; AN=abscess and inflammatory nodule; BHI=BE HEARD I; BHII=BE HEARD II; BKZ=bimekizumab; DT=draining tunnel; HiSCR=Hidradenitis Suppurativa Clinical Response; HiSCR50/75= $\geq 50/75\%$ reduction from baseline in the total AN count with no increase from baseline in abscess or DT count; MI: multiple imputation; mNRI=modified non-responder imputation; OC=observed case; Q2W=every 2 weeks; Q4W=every 4 weeks.

Figure S5. HiSCR90 responses over time to week 48



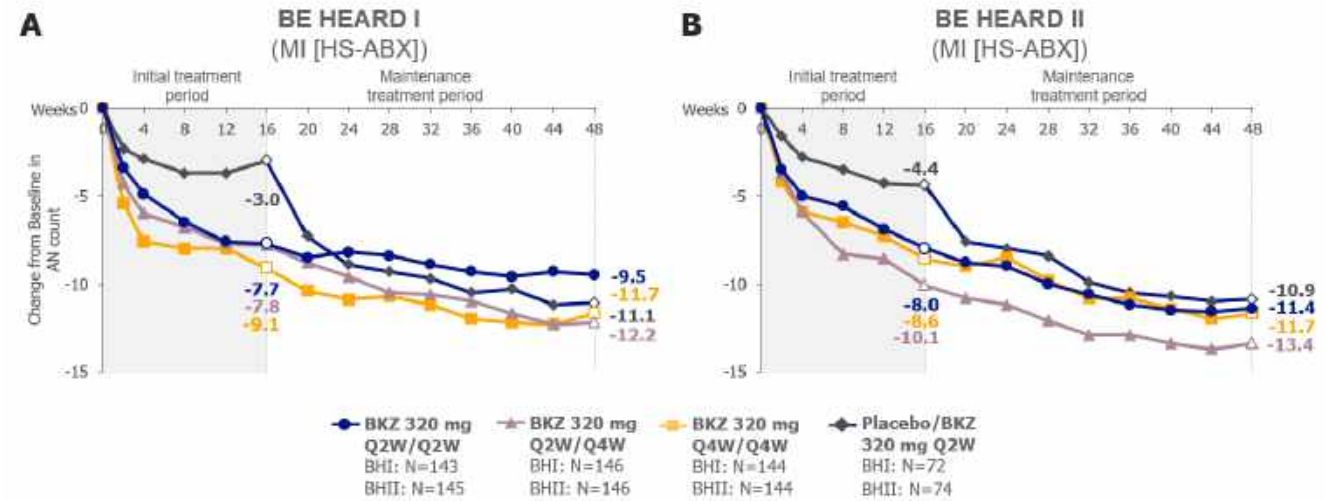
The rates of HiSCR90 over time to week 48 in BE HEARD I and BE HEARD II. mNRI (HS-ABX): patients who took systemic antibiotics defined as rescue medication for HS by the principal investigator or who discontinued due to adverse events or lack of efficacy were treated as non-responders at all subsequent visits. Other missing data were imputed via MI. OC: all available data after an intercurrent event were summarised as recorded in the database, and all missing data were left missing. ABX: antibiotics; AN: abscess and inflammatory nodule; BHI: BE HEARD I; BHII: BE HEARD II; BKZ: bimekizumab; DT: draining tunnel; HiSCR: Hidradenitis Suppurativa Clinical Response; HiSCR90: $\geq 90\%$ reduction from baseline in the total AN count with no increase from baseline in abscess or DT count; HS: hidradenitis suppurativa; MI: multiple imputation; mNRI: modified non-responder imputation; OC: observed case; Q2W: every 2 weeks; Q4W: every 4 weeks.

417 **Figure S6. HiSCR100 responses over time to week 48**



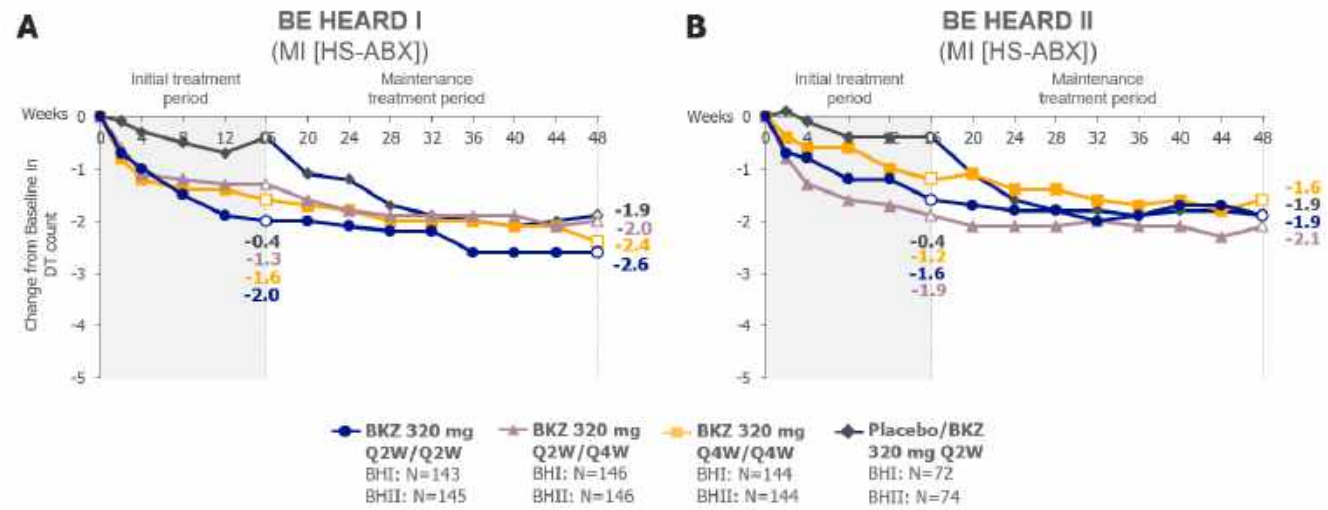
418 The rates of HiSCR100 over time to week 48 in BE HEARD I and BE HEARD II. mNRI (HS-
 419 ABX): patients who took systemic antibiotics defined as rescue medication for HS by the principal
 420 investigator or who discontinued due to adverse events or lack of efficacy were treated as non-
 421 responders at all subsequent visits. Other missing data were imputed via MI. OC: all available data
 422 after an intercurrent event were summarised as recorded in the database, and all missing data were left
 423 missing. ABX: antibiotics; AN: abscess and inflammatory nodule; BHI: BE HEARD I; BHII: BE
 424 HEARD II; BKZ: bimekizumab; DT: draining tunnel; HiSCR: Hidradenitis Suppurativa Clinical
 425 Response; HiSCR100: 100% reduction from baseline in the total AN count with no increase from
 426 baseline in abscess or DT count; HS: hidradenitis suppurativa; MI: multiple imputation;
 427 mNRI: modified non-responder imputation; OC: observed case; Q2W: every 2 weeks; Q4W: every 4
 428 weeks.

Figure S7. Change from baseline in AN count to week 48



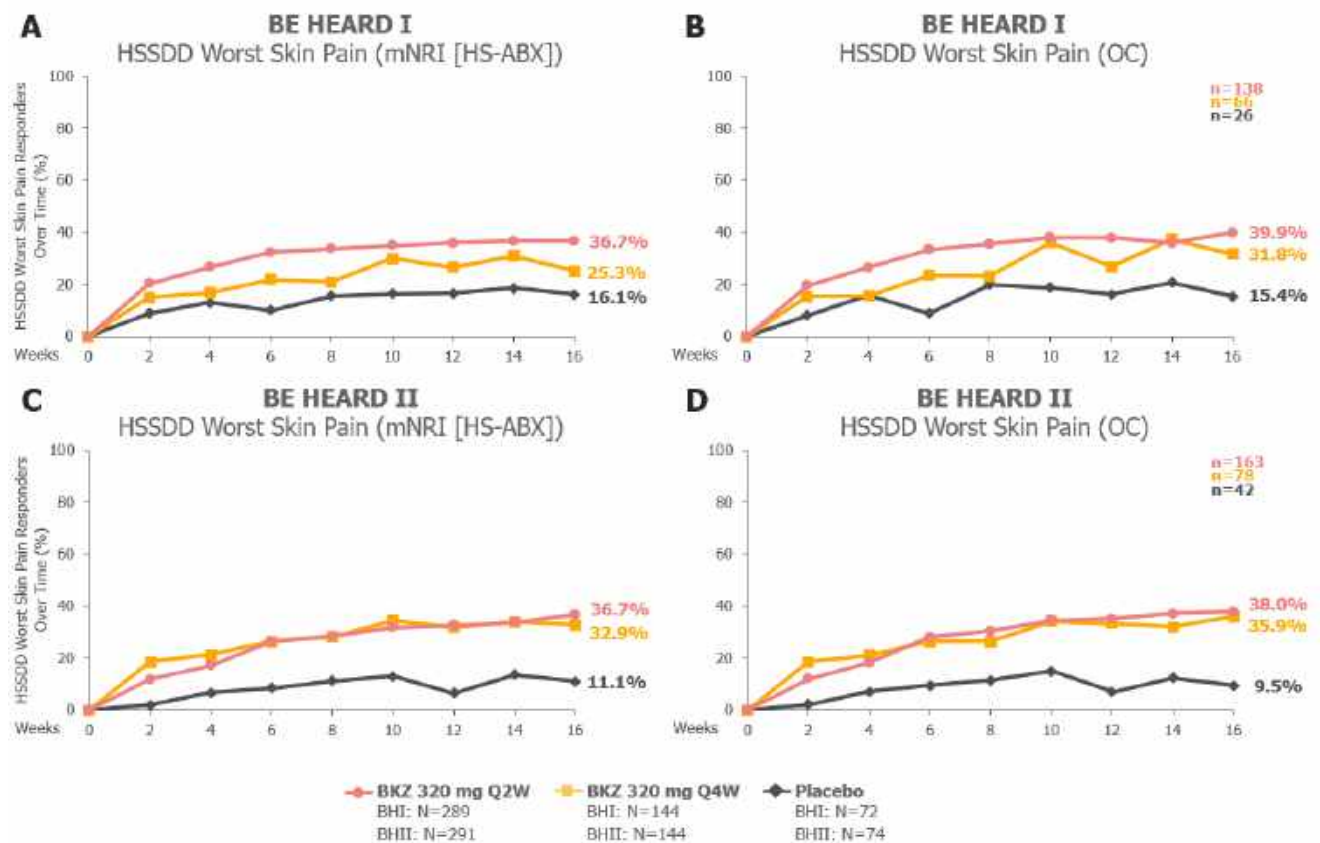
Change from baseline in AN count over time to week 48 in BE HEARD I and BE HEARD II. Data were imputed using MI (HS-ABX): patients who discontinued study treatment due to lack of efficacy or adverse events, or who took systemic antibiotics as rescue medication for HS as defined by the principal investigator, were set to missing and subsequently imputed using MI. ABX: antibiotics; AN: abscess and inflammatory nodule; BHI: BE HEARD I; BHII: BE HEARD II; BKZ: bimekizumab; HS: hidradenitis suppurativa; MI: multiple imputation; Q2W: every 2 weeks; Q4W: every 4 weeks.

Figure S8. Change from baseline in DT count to week 48



Change from baseline in DT count over time to week 48 in BE HEARD I and BE HEARD II. Data were imputed using MI (HS-ABX): patients who discontinued study treatment due to lack of efficacy or adverse events, or who took systemic antibiotics as rescue medication for HS as defined by the principal investigator, were set to missing and subsequently imputed using MI. ABX: antibiotics; BHI: BE HEARD I; BHII: BE HEARD II; BKZ: bimekizumab; DT: draining tunnel; HS: hidradenitis suppurativa; MI: multiple imputation; Q2W: every 2 weeks; Q4W: every 4 weeks.

448 **Figure S9. HSSDD Worst Skin Pain Response over time to week 16**



449 The rates of HSSDD worst skin pain response over time to week 16 in BE HEARD I and BE HEARD
 450 II. mNRI (HS-ABX): patients who took systemic antibiotics defined as rescue medication for HS by
 451 the principal investigator or who discontinued due to adverse events or lack of efficacy were treated as
 452 non-responders at all subsequent visits. Other missing data were imputed via MI. OC: all available data
 453 after an intercurrent event were summarised as recorded in the database, and all missing data were left
 454 missing. Pain response was defined as an improvement from baseline in HSSDD weekly worst skin
 455 pain score of at least 3 points among patients with a baseline score of 3 or higher. ABX: antibiotics;
 456 BHI: BE HEARD I; BHII: BE HEARD II; BKZ: bimekizumab; HS: hidradenitis suppurativa;
 457 HSSDD: HS Symptom Daily Diary; MI: multiple imputation; mNRI: modified non-responder
 458 imputation; OC: observed case; Q2W: every 2 weeks; Q4W: every 4 weeks.

THE LANCET

Supplementary appendix 2

This appendix formed part of the original submission and has been peer reviewed.
We post it as supplied by the authors.

Supplement to: Kimball AB, Jemec GBE, Sayed CJ, et al. Efficacy and safety of bimekizumab in patients with moderate-to-severe hidradenitis suppurativa (BE HEARD I and BE HEARD II): two 48-week, randomised, double-blind, placebo-controlled, multicentre phase 3 trials. *Lancet* 2024; published online May 22. [https://doi.org/10.1016/S0140-6736\(24\)00101-6](https://doi.org/10.1016/S0140-6736(24)00101-6).

SUPPLEMENTARY APPENDIX 2

This appendix has been provided by the authors to give readers additional information about their work.

Table of Contents

BE HEARD I Study Protocol Amendment 5	Page 2
BE HEARD II Study Protocol Amendment 4	Page 126
BE HEARD I SAP Amendment 3	Page 248
BE HEARD II SAP Amendment 2	Page 465

**A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-
CONTROLLED, MULTICENTER STUDY EVALUATING THE
EFFICACY AND SAFETY OF BIMEKIZUMAB IN STUDY
PARTICIPANTS WITH MODERATE TO SEVERE
HIDRADENITIS SUPPURATIVA**

PROTOCOL HS0003 AMENDMENT 5

PHASE 3

SHORT TITLE:

A Phase 3, Randomized, Double-blind, Placebo-Controlled, Multicenter Study Evaluating the Efficacy and Safety of Bimekizumab in Study Participants with Moderate to Severe Hidradenitis Suppurativa

Sponsor:

UCB Biopharma SRL

Allée de la Recherche 60

1070 Brussels

BELGIUM

Regulatory agency identifying numbers:

Eudra CT Number:	2019-002550-23
Investigational New Drug (IND) Number:	134700
NCT Number:	NCT04242446

Confidential Material

Confidential

This document is the property of UCB and may not – in full or in part – be passed on, reproduced, published, or otherwise used without the express permission of UCB.

PROTOCOL AMENDMENT SUMMARY OF CHANGES TABLE

Document History		
Document	Date of Issue	Type of Amendment
Amendment 5	27 Sep 2022	Substantial
Amendment 4	09 May 2022	Substantial
Amendment 3	03 Feb 2021	Substantial
Amendment 2	16 Dec 2019	Nonsubstantial
Amendment 1	06 Dec 2019	Nonsubstantial
Original Protocol	29 Oct 2019	Not applicable

SERIOUS ADVERSE EVENT REPORTING

Serious adverse event reporting (24h)	
Fax	Europe and Rest of the World: +32 2 386 24 21 US and Canada: +1 800 880 6949 or +1 866 890 3175
Email	Global: DS_ICT@ucb.com (for interventional clinical studies)

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Amendment 5 (27 Sep 2022)

Overall rationale for the amendment

Periodic worsening of HS symptoms is a well-known aspect of HS. This worsening has been described as “flaring” by both patients and physicians (Caposiena and Bianchi, 2020). However, there are several different definitions of flare, and no unanimous consensus has evolved over the course of the study on its meaning, including the flare definition used in previous and ongoing clinical studies (Kimball, 2016c and CT.gov citations for HS0003 and HS0004 studies). With no published validation studies, no information is available on the validity, repeatability, or sensitivity to change of any of the described definitions (Kirby 2020). Highlighting these points, inconsistent data on the flare endpoint have been observed within the bimekizumab program and in recently published data from another experimental HS treatment (Kimball, 2022). Given this information, which is external to and independent of any knowledge of the results of this study, the relevance of the flare endpoint as currently defined is questionable. For these reasons, the protocol is amended to remove the flare (secondary) endpoint from the statistical testing procedure. It will be included as an “Other” endpoint in the study.

Section # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3 Objectives and endpoints	Removed flare from secondary objectives	To align with the overall rationale for the amendment
3 Objectives and endpoints	Added flare to “Other” objectives	
9.3.2 Analysis of secondary efficacy endpoints; Figure 9.1; Table 9.1	Removed flare from testing procedures	
References	Added references to support Protocol Amendment 5 overall rationale	
Throughout	Minor editorial and formatting revisions	Minor edits and formatting revisions that do not impact content were made for readability and/or clarity

TABLE OF CONTENTS

1	PROTOCOL SUMMARY	10
1.1	Synopsis	10
1.2	Schema	13
1.3	Schedule of activities	14
2	INTRODUCTION	20
2.1	Study rationale	20
2.2	Background	20
2.3	Benefit/risk assessment	21
3	OBJECTIVES AND ENDPOINTS	23
4	STUDY DESIGN	26
4.1	Overall design	26
4.1.1	Screening Period (Weeks -5 to 0)	27
4.1.2	Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)	27
4.1.3	Safety Follow-up Visit	27
4.1.4	Visit Windows	27
4.2	Scientific rationale for study design	27
4.3	Justification for dose	28
4.4	End of study definition	29
5	STUDY POPULATION	29
5.1	Inclusion criteria	29
5.2	Exclusion criteria	30
5.3	Lifestyle restrictions	34
5.4	Screen failures	34
6	STUDY TREATMENTS	35
6.1	Treatments administered	35
6.2	Preparation, handling, storage, and accountability requirements	38
6.2.1	Drug accountability	38
6.3	Measures to minimize bias: randomization and blinding	38
6.3.1	Randomization and numbering of participants	38
6.3.2	Procedures for maintaining and breaking the treatment blind	39
6.3.2.1	Maintenance of study treatment blind	39
6.3.2.2	Breaking the treatment blind in an emergency situation	39
6.4	Treatment compliance	40
6.5	Concomitant medications/treatments	40
6.5.1	Permitted concomitant treatments (medications and therapies)	40
6.5.2	Prohibited concomitant treatments (medications and therapies)	41

6.5.3	Rescue medication	44
6.5.3.1	Antibiotic Rescue Medication/Antibiotics	44
6.5.3.2	Rescue Medications/Lesion Intervention	44
6.6	Dose modification	45
6.7	Criteria for study hold or dosing stoppage.....	45
6.8	Treatment after the end of the study	45
7	DISCONTINUATION OF STUDY MEDICATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL.....	46
7.1	Discontinuation of study medication	46
7.1.1	Study participant does not achieve partial response	46
7.1.2	Potential drug-induced liver injury IMP discontinuation criteria.....	46
7.1.2.1	PDILI Discontinuation Criteria.....	46
7.1.3	Treatment interruptions	47
7.2	Participant discontinuation/withdrawal from the study	48
7.3	Lost to follow up.....	50
8	STUDY ASSESSMENTS AND PROCEDURES.....	51
8.1	Efficacy assessments.....	52
8.1.1	Lesion count.....	52
8.1.1.1	Hidradenitis Suppurativa Clinical Response (HiSCR).....	52
8.1.1.2	Hidradenitis Suppurativa Physician's Global Assessment.....	53
8.1.1.3	International Hidradenitis Suppurativa Severity Score System (IHS4)	53
8.1.2	Partial response	53
8.1.3	High-sensitivity C-reactive protein (hs-CRP)	54
8.1.4	Patient-reported outcomes	54
8.1.4.1	HS symptom measures of skin pain, smell or odor, drainage or oozing from HS lesions, and itch	54
8.1.4.2	Patient Global Impression of HS Severity (PGI-S-HS) and Change in HS Severity (PGI-C-HS)	54
8.1.4.3	Patient Global Impression of Severity of Skin Pain (PGI-S-SP) and Change in Severity of Skin Pain (PGI-C-SP).....	54
8.1.4.4	Hidradenitis Suppurativa Quality of Life Questionnaire (HiSQOL)	55
8.1.4.5	Dermatology Life Quality Index (DLQI)	55
8.1.4.6	Euro-Quality of Life 5-Dimensions, 3 levels	55
8.1.4.7	Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem.....	55
8.1.4.8	Treatment Satisfaction Questionnaire for Medication	55
8.1.5	Hurley Stage	56
8.2	Safety assessments.....	56
8.2.1	Physical examination	56

8.2.2	Vital signs	56
8.2.3	Electrocardiograms	57
8.2.4	Clinical safety laboratory assessments	57
8.2.5	Depression and suicidal risk monitoring	57
8.2.5.1	PHQ-9.....	57
8.2.5.2	eC-SSRS	57
8.2.6	Assessment and management of tuberculosis and tuberculosis risk factors.....	58
8.2.6.1	Assessments at Screening.....	58
8.2.6.2	Definitions	58
8.2.6.3	Assessment and reporting of TB and TB risk factors during the study.....	59
8.3	Adverse events and serious adverse events	64
8.3.1	Time period and frequency for collecting AE and SAE information.....	64
8.3.2	Method of detecting AEs and SAEs	64
8.3.3	Follow-up of AEs and SAEs.....	64
8.3.4	Regulatory reporting requirements for SAEs	65
8.3.5	Pregnancy	65
8.3.6	Adverse events of special interest.....	65
8.3.7	Other safety topics of interest	66
8.3.8	Anticipated serious adverse events	66
8.3.9	Suspected transmission of an infectious agent.....	67
8.4	Safety signal detection	67
8.5	Treatment of overdose	68
8.6	Pharmacokinetics	68
8.7	Genetics.....	68
8.8	Pharmacodynamics	69
8.9	Biomarkers	69
8.9.1	Immunogenicity assessments.....	69
8.10	Medical resource utilization and health economics	70
8.11	Photography	70
9	STATISTICAL CONSIDERATIONS.....	70
9.1	Definition of analysis sets.....	70
9.1.1	Enrolled Set	70
9.1.2	Randomized Set	70
9.1.3	Safety Set	70
9.1.4	Full Analysis Set.....	70
9.1.5	Per-Protocol Set	70
9.1.6	Pharmacokinetics Per-Protocol Set.....	71
9.1.7	COVID-19 Free Set	71

9.2	General statistical considerations.....	71
9.3	Planned efficacy/outcome analyses	71
9.3.1	Analysis of the primary efficacy endpoint.....	71
9.3.2	Analysis of the secondary efficacy endpoints	72
9.3.3	Other efficacy/other outcome analyses.....	74
9.3.4	COVID-19 impact analysis.....	74
9.4	Planned safety and other analyses.....	75
9.4.1	Safety analyses.....	75
9.4.2	PK and ADAb analyses	75
9.5	Handling of protocol deviations.....	75
9.6	Handling of dropouts or missing data.....	75
9.7	Planned interim analysis and data monitoring	77
9.7.1	Data Monitoring Committee.....	77
9.7.2	Interim Analysis.....	77
9.8	Determination of sample size.....	77
10	SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS... 80	
10.1	Appendix 1: Regulatory, ethical, and study oversight considerations.....	80
10.1.1	Regulatory and ethical considerations	80
10.1.2	Financial disclosure	80
10.1.3	Informed consent process.....	81
10.1.4	Data protection.....	81
10.1.5	Committees structure.....	82
10.1.6	Data quality assurance	82
10.1.6.1	Case report form completion.....	83
10.1.7	Source documents.....	83
10.1.8	Study and Site Closure.....	83
10.1.9	Publication Policy	84
10.2	Appendix 2: Clinical laboratory tests	85
10.3	Appendix 3: Adverse events – definitions and procedures for recording, evaluating, follow-up, and reporting	87
10.4	Appendix 4: Contraceptive guidance and collection of pregnancy information	93
10.5	Appendix 5: Genetics.....	96
10.6	Appendix 6: Liver safety – suggested actions and follow-up assessments	97
10.6.1	Consultation with Medical Monitor and local hepatologist.....	101
10.6.2	Immediate action: determination of IMP discontinuation	101
10.6.2.1	IMP restart/rechallenge	101
10.6.3	Testing: identification/exclusion of alternative etiology	102

10.7	Appendix 7: Medical device AEs, Adverse device effects, SAEs, and device deficiencies: definition and procedures for recording, evaluating, follow-up, and reporting.....	105
10.8	Appendix 8: Rapid alert procedures	106
10.9	Appendix 9: Country-specific requirements.....	107
10.10	Appendix 10: Abbreviations and trademarks	108
10.11	Appendix 11: Protocol amendment history	111
11	REFERENCES	120
	SPONSOR DECLARATION	123

LIST OF TABLES

Table 1–1:	Schedule of activities	14
Table 6–1:	Study medications administered	36
Table 6–2:	Dosing scheme	37
Table 6–3:	Exclusions for prior medications	42
Table 8–1:	Anticipated SAEs for the Population of Participants with HS	67
Table 9–1:	Power calculation assumptions and methods.....	79
Table 10–1:	Required investigations and follow-up for PDILI	98
Table 10–2:	PDILI laboratory measurements	103
Table 10–3:	PDILI information to be collected	104

LIST OF FIGURES

Figure 1-1:	Study Schematic.....	13
Figure 8-1:	Decision tree for IGRA TB results at Screening	62
Figure 8-2:	Decision tree for IGRA TB results during a study	63
Figure 9–1:	Closed Testing Procedure	74

1 PROTOCOL SUMMARY

1.1 Synopsis

Protocol title:

A Phase 3, Randomized, Double-blind, Placebo-Controlled, Multicenter Study Evaluating the Efficacy and Safety of Bimekizumab in Study Participants with Moderate to Severe Hidradenitis Suppurativa

Short Title:

A Phase 3, Randomized, Double-blind, Placebo-Controlled, Multicenter Study Evaluating the Efficacy and Safety of Bimekizumab in Study Participants with Moderate to Severe Hidradenitis Suppurativa

Rationale:

UCB is investigating bimekizumab (a humanized, full-length immunoglobulin [Ig] G1 monoclonal antibody [mAb] with high affinity for human interleukin [IL] 17 [IL-17A and IL-17F]) for the treatment of hidradenitis suppurativa (HS). Antibodies targeting IL-17A are effective in treating patients with moderate to severe psoriasis (PSO) and other immuno-inflammatory conditions. Data from clinical studies with bimekizumab suggest that inhibition of both IL-17A and IL-17F could be beneficial to patients with such conditions, including HS. Based on results from a Phase 2a study in study participants with HS (HS0001); and a Phase 2b study in study participants with PSO (PS0010) and its 48 week extension study (PS0011), bimekizumab doses of 320mg every 2 weeks (Q2W) in HS and 320mg every 4 weeks (Q4W) in PSO appeared to have an acceptable safety profile, and achieved clinically meaningful efficacy in both primary and key secondary endpoints in their respective studies. Based on these data and considering the high unmet need for safe and effective therapies for HS, confirmatory studies are being conducted with bimekizumab for the treatment of moderate to severe HS.

Objectives and Endpoints

The primary and secondary objectives and associated endpoints are as follows:

Objectives	Endpoints
Primary	
Evaluate the efficacy of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₅₀ at Week 16
Secondary	
Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₇₅ at Week 16 Absolute change from Baseline in DLQI Total Score at Week 16 Absolute change from Baseline in the Worst HS Skin Pain score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HSSDD Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) at Week 16
Evaluate the safety of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> Treatment-emergent AEs Serious TEAEs TEAEs leading to withdrawal from study

AE=adverse event; AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS= hidradenitis suppurativa; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; TEAE=treatment-emergent adverse event

Overall Design

HS0003 is a Phase 3, randomized, double-blind, placebo-controlled, multicenter, pivotal study evaluating the efficacy and safety of bimekizumab in study participants with moderate to severe HS.

Number of Participants

Approximately 490 study participants will be randomly assigned to study treatment: 140 for bimekizumab 320mg Q2W/Q2W, 140 for bimekizumab 320mg Q2W/Q4W, 140 for bimekizumab 320mg Q4W/Q4W, and 70 for placebo/320mg Q2W. The Randomized Set (ie, all randomized study participants) is the primary analysis set for efficacy analyses.

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomized to study treatment (Section 5.4).

Treatment Groups and Duration

Study participants meeting the inclusion criteria who do not meet any exclusion criteria will complete a Screening Period of 14 days to up to 5 weeks; a double-blind, 48-week Treatment Period comprising a 16-week Initial Treatment Period and 32-week Maintenance Treatment Period; and a 20-week Safety Follow-up (SFU) Period following the final injection of investigational medicinal product (IMP) if study participants do not enter a subsequent extension study (HS0005) or withdraw prematurely from treatment. The total duration of study participation in HS0003 will be 68 to 71 weeks for those who complete HS0003 and do not participate in the extension study HS0005 and 50 to 53 weeks for those who participate in HS0005 and, thus, do not participate in the 20-week SFU Period.

Study participants will be randomized in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to 1 of 3 dose regimens of bimekizumab or placebo as shown in the schematic (Figure 1-1). All doses of IMP will be administered by subcutaneous (sc) injection. The primary efficacy variable at Week 16 is HiSCR₅₀ (a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count). Study visits will occur at Screening; Baseline (Week 0); Weeks 1, 2, 4, 6, 8, 10, 12, 14, and 16; every 2 weeks from Week 16 through Week 48 for assessments of efficacy, safety, and other measures of quality of life/health status/work productivity; and a SFU visit 20 weeks after the last dose of IMP for participants who do not enter the extension study.

An independent Data Monitoring Committee (DMC) will periodically review unblinded efficacy and safety data to assess the benefit/risk of bimekizumab in study participants with moderate to severe HS.

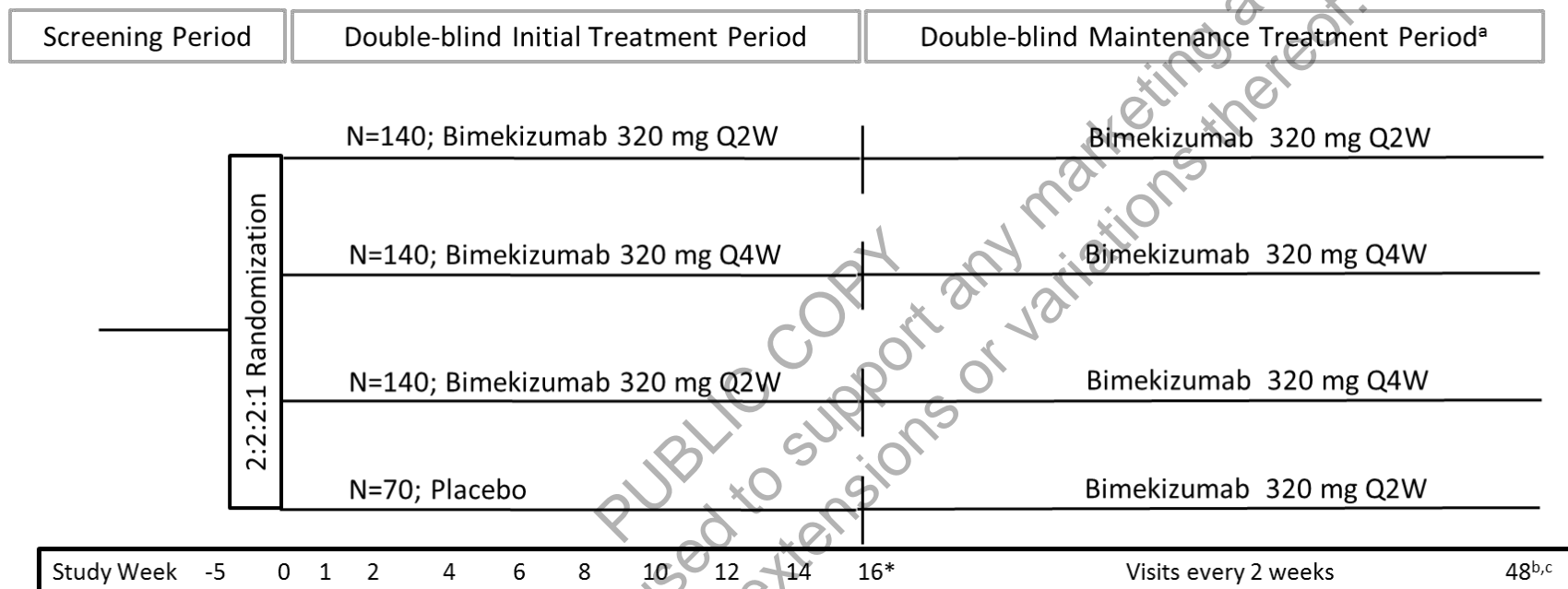
[REDACTED]

[REDACTED]

1.2 Schema

A schematic diagram for the study is presented in [Figure 1-1](#).

Figure 1-1: Study Schematic



HiSCR₅₀=a 50% reduction in the total abscess and inflammatory nodule count with no increase from Baseline in abscess or draining tunnel count;

IMP=investigational medicinal product; Q2W=every 2 weeks; Q4W=every 4 weeks

*Week 16 = primary endpoint (HiSCR₅₀ bimekizumab versus placebo)

^a Study participants should discontinue from the study from Week 32 on if no partial response is achieved (partial response is defined as $\geq 25\%$ improvement in abscess and inflammatory nodule count relative to Baseline [Week 0] lesion values.)

^b Study participants achieving an improvement of at least 25 % in abscess and inflammatory nodule count continue in HS0005 (Extension Study).

^c 20-week Safety Follow-up (from last IMP injection) for any study participant who discontinues from study prior to Week 48, or who does not continue in HS0005.

1.3 Schedule of activities

The Schedule of Activities is presented in [Table 1–1](#).

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)											Maintenance Treatment Period (weeks after first dose)															PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Study Visit	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Informed consent	X																										X ^d		
Study participant number assigned	X																												
Inclusion/exclusion	X	X																											
Demographic and Baseline disease characteristics	X	X																											
Hidradenitis suppurativa history	X																												
Significant past medical history/concomitant diseases ^e	X	X ^f																											
Physical examination ^g	X	X							X		X				X						X						X	X	X

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Height		X																											
Body weight		X													X												X	X	
Vital signs ^h	X	X	X	X	X		X		X		X	X	X		X		X		X		X		X		X		X	X	X
12-lead ECG	X										X								X								X	X	X
Hematology/ biochemistry	X	X		X	X		X		X		X	X	X		X		X		X		X		X		X		X	X	X
hs-CRP		X							X		X				X				X		X						X	X	
Urinalysis	X	X									X				X				X				X				X	X	X
Urine drug screen	X																										X	X	
Pregnancy testing ⁱ	X	X			X		X		X		X		X		X		X		X		X		X		X		X	X	X
Hepatitis B and C testing ^j	X																												
HIV testing ^k	X																												
Chest x-ray ^l	X																												
IGRA TB test ^m	X																								X				
TB questionnaire	X	X							X						X						X						X	X	X
Lesion count (includes Hurley Stage) ⁿ	X	X		X	X		X		X		X		X		X		X		X		X		X		X		X	X	

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Photography ^o		X			X						X								X								X	X	
Blood sample – bimekizumab plasma concentrations ^p		X	X	X	X		X		X		X	X	X	X	X						X						X	X	X
Blood sample for anti-drug antibodies ^p		X			X		X		X		X	X	X	X	X						X						X	X	X
Blood sample for genomic/ proteomic/ metabolomics, and candidate biomarker analyses ^{p,q}		X					X		X		X																X	X	
Blood sample for genetic/ epigenetic analyses ^{p,q}		X									X																X	X	
Urine samples for biomarker research		X									X																X	X	
DLQI		X			X		X		X		X		X						X		X						X	X	
PHQ-9	X	X			X		X		X		X		X		X		X		X		X		X		X		X	X	X

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFT ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
eC-SSRS	X	X			X		X		X		X		X		X		X		X		X		X		X		X	X	X
WPAI-SHP		X							X		X								X								X	X	
TSQM-9											X															X	X		
HiSQOL		X			X						X								X							X	X		
HSSDD ^r	●										●																		
HSSQ		X									X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PGI-S-HS		X			X						X								X							X	X		
PGI-C-HS					X						X								X							X	X		
PGI-S-SP		X			X						X								X							X	X		
PGI-C-SP					X						X								X							X	X		
EQ-5D-3L		X			X				X		X								X		X					X	X		
Concomitant medications	●																											●	
Adverse events	●																											●	
Randomization ^s		X																											
IRT ^t	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Bimekizumab/ placebo administration ^t		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X				

AN=abscess and inflammatory nodule; Bsl=Baseline; CAT=computed axial tomography; CRF=case report form; DLQI=Dermatology Life Quality Index; ECG=electrocardiogram; eC-SSRS=electronic Columbia-Suicide Severity Rating Scale; EQ-5D-3L=Euro-Quality of Life 5-Dimensions, 3 levels; HiSCR₂₅=a 25% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₅₀=a 50% reduction in the total

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Protocol activity																													

AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₉₀=a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₁₀₀=a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSQOL=Hidradenitis Suppurativa Quality of Life; HIV=human immunodeficiency virus; HS=hidradenitis suppurativa; hs-CRP=high sensitivity C-reactive protein; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; HSSQ=Hidradenitis Suppurativa Symptom Questionnaire; IGRA=interferon-gamma release assay; IMP=investigational medicinal product; IRT=interactive response technology; PEOT=Premature End of Treatment Visit; PGI-C-HS=Patient Global Impression of Change in HS Severity; PGI-C-SP=Patient Global Impression of Change in Severity of Skin Pain; PGI-S-HS=Patient Global Impression of HS Severity; PGI-S-SP=Patient Global Impression of Severity of Skin Pain; PHQ-9=Patient Health Questionnaire 9; SFU=Safety Follow-Up; TB=tuberculosis; TSQM=Treatment Satisfaction Questionnaire for Medication; WPAI-SHP=Work Productivity and Impairment Questionnaire-Specific Health Problem

Note: Some study-specific assessments may be conducted remotely according to the study protocol during a pandemic or other exceptional circumstance (eg, hurricanes) due to the need to implement safety measures and guidance from regulatory authorities. Refer to Section 8 for additional information.

^a Visit windows are ± 3 days (based on the date of the first dose). However, the minimum of 8 days between injection visits (eg, Visit 2 +3 days occurs on Day17; Visit 4 -3 days occurs on Day 25) may be used only 1 time in the Initial Treatment Period and 1 time in the Maintenance Treatment Period, if needed. The study participant should be dosed according to the administration schedule thereafter. The 20-week SFU Visit window is ± 7 days (based on the date of the final dose).

^b The Screening Period will last a minimum of 14 days up to a maximum of 5 weeks prior to randomization.

^c All study participants not continuing in the extension study, including those withdrawn from IMP as well as those completing all study visits, will have an SFU Visit 20 weeks after his or her final dose of IMP.

^d A separate informed consent form will be required to be completed for participation in extension study HS0005. Study participants continuing in HS0005 will receive first dose of bimekizumab in that study on Week 48.

^e Past medical history includes tobacco and alcohol use.

^f Ensure no significant changes from Baseline in medical history and concomitant disease.

^g The physical examination will be performed as detailed in Section 8.2.1. Includes evaluation of signs and symptoms of active TB and risk for exposure to TB.

^h Vital signs (sitting systolic and diastolic blood pressure, pulse rate, and body temperature) are to be measured prior to blood sampling and prior to dosing, where applicable.

ⁱ Pregnancy testing will consist of serum testing at the Screening Visit. Urine pregnancy tests will be performed at all other visits where specified.

^j See Exclusion Criterion 9.

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Protocol activity																													

^k The HIV test result will not be recorded in the eCRF.

^l A chest x-ray must be performed at the Screening Visit, or must occur within 2 months prior to Screening and results must be available at Baseline. A CAT scan of chest at Screening or within 2 months prior to Screening is acceptable, if available. The official report of the imaging must specifically indicate the absence of findings suggestive of current active TB or prior inactive TB.

^m Quantiferon Gold Plus will be used for this analysis. Details on IGRA assessment are provided in Section 8.2.6.3.4.

ⁿ Lesion count will be performed at the specified study visits, and must address all relevant anatomical regions in each study participant; Hurley Stage is included as an assessment performed during the lesion count. The data collected from the lesion count will be used for the derivation of the HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀, HS Physician's Global Assessment, and International Hidradenitis Suppurativa Severity score system (IHS4).

^o At certain sites, where feasible, representative photographs of the changes in skin will be captured. Photographs will be anonymized.

^p All blood samples taken prior to dosing.

^q All genomic, proteomic, and metabolomic samples and the genetic/epigenetic samples will be stored at -80°C at the central biorepository for up to 20 years.

^r The HSSDD (pain, smell or odor, drainage or oozing from HS lesions, and itch) will be completed daily from Screening through Week 16.

^s Randomization occurs at the Baseline Visit for all study participants.

^t IMP administration is based on randomization at the Baseline Visit for all study participants. Double-dummy IMP administration will occur every 2 weeks in all study participants to maintain the study blind. The dosing window is ± 3 days relative to the scheduled dosing visit.

2 INTRODUCTION

2.1 Study rationale

UCB is investigating bimekizumab (a humanized, full-length immunoglobulin [Ig] G1 monoclonal antibody [mAb] with high affinity for human interleukin [IL] 17 [IL-17A and IL 17F]) for the treatment of hidradenitis suppurativa (HS). Antibodies targeting IL-17A are effective in treating patients with moderate to severe psoriasis (PSO) and other immuno-inflammatory conditions. Data from clinical studies with bimekizumab suggest that inhibition of both IL-17A and IL-17F could be beneficial to patients with such conditions, including HS. Based on results from a Phase 2a study in study participants with HS (HS0001; European Union Drug Regulating Authorities Clinical Trials [EudraCT] Number 2017-000892-10, NCT03248531) and a Phase 2b study in study participants with PSO (PS0010) and its 48 week extension study (PS0011), bimekizumab doses of 320mg every 2 weeks (Q2W) in HS and 320mg every 4weeks (Q4W) in PSO, appeared to have an acceptable safety profile, and achieved clinically meaningful efficacy in both primary and key secondary endpoints in their respective studies. Based on these data and considering the high unmet need for safe and effective therapies for HS, confirmatory studies are being conducted with bimekizumab for the treatment of moderate to severe HS.

2.2 Background

Hidradenitis suppurativa or acne inversa is a chronic, inflammatory, recurrent, debilitating skin disease that usually presents after puberty with painful, deep-seated, inflamed lesions in the apocrine gland-bearing areas of the body, most commonly the axillaries, inguinal, and anogenital regions (Dessau definition, First International Conference on HS, 30 Mar to 01 Apr 2006, Dessau, Germany). The nodules are often inflamed, can progress to abscess formation, and may rupture to form fistulas and subsequent scarring. Thus, many patients with HS develop permanent sequelae of past inflammation that are only remediable through surgical excision of the involved skin areas. Hidradenitis suppurativa is also associated with several complications (eg, the development of anal, urethral, and rectal strictures and fistulas), and the excessive scarring and fibrosis produced by HS lesions can lead to contractures and limitations in limb mobility (Alikhan, 2009; Danby, 2010).

Hidradenitis suppurativa is estimated to affect about 1% of the adult European population, with a female to male ratio of approximately 3:1 (Revuz, 2008; Naldi, 2006). The prevalence of diagnosed HS in the US may be lower (<0.1%), although further research is needed to determine the prevalence of undiagnosed HS in the US (McMillan, 2014). Patients diagnosed with HS often experience a significant reduction in quality of life (QOL), equivalent to severe PSO (Sartorius, 2009), due to the location of, and discharge from, the lesions that leads to an often persistent morbidity due to pain and sequelae from uncontrolled inflammation (von der Werth, 2001; Wolkenstein, 2007). The reduction in QOL and persistent morbidity result in functional impairment in patients with HS similar or greater to that of heart disease, diabetes, or asthma, when measured by the European Quality-of-Life 5 dimensions 3-level questionnaire (EQ-5D-3L) scale (Riis, 2016).

Bimekizumab is a humanized full-length mAb of IgG1 subclass being developed for the treatment of patients with inflammatory diseases such as PSO, psoriatic arthritis, axial spondyloarthritis, and HS. Bimekizumab has high affinity for human IL-17A and human IL-17F,

and selectively and potently inhibits the activity of both isoforms in vitro. The key pro-inflammatory cytokine IL-17A has been demonstrated to, and IL-17F is believed to, play important roles in autoimmune and inflammatory diseases. Published data and immunohistochemistry studies performed by UCB have shown that expression of both IL-17A and IL-17F is present in HS lesions, and there are published reports highlighting the potential for IL-17A and IL-17F to drive HS disease pathology (UCB Research Report 40001864; [Cho, 2012](#); [Schlapbach, 2011](#)). This supports the hypothesis that the IL-17 cytokine family is a potential therapeutic target in HS. Bimekizumab is hypothesized to demonstrate a treatment response in HS because it selectively and potently inhibits the activity of both IL-17A and IL-17F isoforms in vitro.

2.3 Benefit/risk assessment

Overall, the Phase 2a POC study in HS (HS0001) demonstrated that the safety profile (during 10 weeks of dosing for a 12-week treatment period) for bimekizumab 320mg sc Q2W appears consistent with that of bimekizumab in other indications (PSO, psoriatic arthritis, axial spondyloarthritis) for which bimekizumab is being developed.

Treatment-emergent adverse events (TEAEs) experienced by study participants with HS receiving repeated doses of bimekizumab occurred at incidences similar to those with placebo and adalimumab (range: 61.9% to 71.4%). In both the bimekizumab and adalimumab groups, the most frequently reported TEAEs were in the system organ classes of Infections and infestations (43.5% and 42.9%, respectively), Skin and subcutaneous tissue disorders (28.3% and 42.9%, respectively), and General disorders and administration site conditions (21.7% and 23.8%, respectively). The most frequently reported TEAEs in the placebo group were in the SOC of Nervous system disorders (28.6%) and Infections and infestations and Skin and subcutaneous tissue disorders (19.0% each). The most frequently reported TEAEs (by preferred term) were hidradenitis (17.4%) and fatigue (8.7%) in the bimekizumab group, hidradenitis (33.3%) and influenza (14.3%) in the adalimumab group, and headache and hidradenitis (14.3% each) and arthralgia (4.8%) in the placebo group. Events of hidradenitis were not unexpected in a population of study participants with moderate-to-severe HS.

With regard to TEAEs of special interest and other safety topics of interest, there were no major adverse cardiovascular events, serious fungal/opportunistic infections (including tuberculosis [TB]), malignancies (including lymphoma), neuropsychiatric events, cases of inflammatory bowel disease (IBD), evidence of hepatotoxicity (per Hy's Law criteria), or hypersensitivity/anaphylactic reactions reported with bimekizumab treatment. No new safety signals were identified in HS0001 compared with those observed with bimekizumab across other development programs to date.

No clinically relevant patterns of changes were observed in any treatment group in hematology, clinical chemistry, vital signs, or electrocardiogram (ECG) findings. Few markedly abnormal post-Baseline liver function test values were reported during the study.

With respect to benefit, the totality of the data from HS0001 demonstrated that bimekizumab results in clinically meaningful and consistent improvements versus placebo across all HS outcome measures assessed. Improvements began early after initiation of treatment, and persisted through the last assessment (Week 12). The efficacy data for the primary endpoint, HiSCR₅₀ (ie, a 50% reduction in the total abscess and inflammatory nodule [AN] count with no

increase from Baseline in abscess or draining tunnel count) at Week 12, were comparable with that of adalimumab, and other efficacy measures suggested improved therapeutic outcomes (larger proportions of study participants achieving HiSCR₇₅ [a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count], HiSCR₉₀ [a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count], and Dermatology Life Quality Index [DLQI] total scores of 0 and 1).

Overall, results of HS0001 show that bimekizumab appears to have an acceptable safety profile considering the anticipated benefit at a dose of 320mg Q2W for the treatment duration evaluated to date. No new safety concerns were raised in study participants with moderate to severe HS, and the benefit/risk remains positive and supports continued investigation of bimekizumab at 320mg Q2W.

More detailed information about the known and expected benefits and risks of bimekizumab may be found in the Investigator's Brochure.

3 OBJECTIVES AND ENDPOINTS

Objectives	Endpoints
Primary	
Evaluate the efficacy of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₅₀ at Week 16
Secondary	
Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₇₅ at Week 16 Absolute change from Baseline in DLQI Total Score at Week 16 Absolute change from Baseline in Skin Pain score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HSSDD Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) at Week 16
Evaluate the safety of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> Treatment-emergent AEs Serious TEAEs TEAEs leading to withdrawal from study
Other	
Evaluate the efficacy of bimekizumab on HiSCR, other HS Scores, and other clinical measures of disease activity at various timepoints in study participants with moderate to severe HS	<ul style="list-style-type: none"> Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ Absolute change from Baseline in International Hidradenitis Suppurativa Severity Score System Change from Baseline in the HS-Physician’s Global Assessment 6-point scale Absolute and percentage change from Baseline in hs-CRP Initiation of systemic antibiotic rescue therapy HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ at both Weeks 16 and 48 Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₅₀ during the Maintenance Treatment Period

Objectives	Endpoints
	<ul style="list-style-type: none"> Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period
<p>Evaluate the efficacy of bimekizumab on abscesses, nodules, and draining tunnels at various timepoints in study participants with moderate to severe HS</p>	<ul style="list-style-type: none"> Change and percentage change from Baseline in lesion counts (abscess count, inflammatory nodule count, AN count, and draining tunnel count) AN count of 0, 1, or 2 AN₂₅, AN₅₀, AN₇₅, AN₉₀, AN₁₀₀ Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16 Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48 Time to flare from Weeks 0 to 16 Time to flare from Week 16 to 48
<p>Evaluate the efficacy of bimekizumab on patient-reported outcome measures at various timepoints in study participants with moderate to severe HS</p>	<ul style="list-style-type: none"> Absolute and percentage change (worst and average pain) from Baseline in HS Skin Pain score (11-point numeric rating scale) Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) Pain response (defined as a decrease from Baseline in HSSQ weekly worst skin pain score at or beyond the threshold for clinically meaningful change) Pain response (at least a 30% reduction and at least a 1-unit reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline Pain response (at least a 30% reduction and at least a 1-unit reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline Pain response (at least a 30% reduction and at least a 1-unit reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline Absolute change from Baseline in DLQI Total Score DLQI Total Score of 0 or 1

Objectives	Endpoints
	<ul style="list-style-type: none"> • Minimum clinically important difference (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4) • Absolute change from Baseline in HiSQOL domain scores (symptoms, psychosocial, activities, adaptations) • Absolute change from Baseline in Patient Global Impression of HS Severity • Absolute change from Baseline in Patient Global Impression of Severity of HS Skin Pain • Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor. • Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor • Responses to the EQ-5D-3L, absolute and changes from Baseline in EQ-5D-3L visual analog scale scores • Absolute change from Baseline in Work Productivity and Activity Impairment Questionnaire-Specific Health Problem v2.0 adapted to HS scores • Domain Scores (effectiveness, convenience and global satisfaction) on the Treatment Satisfaction Questionnaire – Medication-9
Evaluate the effect of bimekizumab on other safety measures at various timepoints in study participants with moderate to severe HS	<ul style="list-style-type: none"> • Adverse events of special interest (Hy's Law) • Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behavior, major adverse cardiovascular events, hepatic events and potential drug-induced liver injury (PDILI), malignancies, and inflammatory bowel disease. • Absolute change from Baseline in the PHQ-9 score • Absolute change from Baseline in vital signs • Absolute change from Baseline in clinical laboratory values (chemistry and hematology) • ECG results
Evaluate the pharmacokinetics of bimekizumab in study participants with moderate to severe HS	Plasma bimekizumab concentrations over the study duration

Objectives	Endpoints
Evaluate the immunogenicity of bimekizumab (antidrug antibodies) in study participants with moderate to severe HS	<ul style="list-style-type: none"> Bimekizumab antidrug antibodies Bimekizumab neutralizing antibodies
Exploratory	
Evaluate biomarkers in study participants with moderate to severe HS.	<ul style="list-style-type: none"> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <p>The candidate exploratory biomarkers are the blood or blood derivative (eg, plasma) concentrations of</p> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div>

AE=adverse event; AN=abscess and inflammatory nodule; AN₂₅=a 25% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₅₀=a 50% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₇₅=a 75% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₉₀=a 90% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₁₀₀=a 100% reduction in the total abscess and inflammatory nodule count relative to Baseline; DLQI=Dermatology Life Quality Index; ECG=electrocardiogram; eC-SSRS=electronic Columbia Suicide Severity Rating Scale; EQ-5D-3L=Euro-Quality of Life 5-Dimensions, 3 levels; HiSCR=Hidradenitis Suppurativa Clinical Response; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₉₀=a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₁₀₀=a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSQOL=Hidradenitis Suppurativa Quality of Life; HS=hidradenitis suppurativa; hs-CRP=high-sensitivity C-reactive protein; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; HSSQ=Hidradenitis Suppurativa Symptom Questionnaire; PDILI=potential drug-induced liver injury; PHQ-9=Patient Health Questionnaire Depression Module; QOL=quality of life; TB=tuberculosis; TEAE=treatment-emergent adverse event

4 STUDY DESIGN

4.1 Overall design

HS0003 is a Phase 3, randomized, double-blind, placebo-controlled, multicenter, pivotal study evaluating the efficacy and safety of bimekizumab in study participants with moderate to severe HS. Study participants meeting the inclusion criteria who do not meet any exclusion criteria will complete a Screening Period of 14 days to up to 5 weeks; a double-blind, 48-week Treatment Period comprising a 16-week Initial Treatment Period and 32-week Maintenance Treatment Period; and a 20-week Safety Follow-up (SFU) Period following the final injection of investigational medicinal product (IMP) if study participants do not enter a subsequent extension study (HS0005) or withdraw prematurely from treatment.

Study participants will be randomized in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to 1 of 3 dose regimens of bimekizumab or placebo as shown in the schematic (Figure 1-1). All doses of IMP will be administered by sc injection. The primary efficacy variable at Week 16 is HiSCR₅₀. Study visits will occur at Screening; Baseline (Week 0); Weeks 1, 2, 4, 6, 8, 10, 12, 14, and 16; and every 2 weeks from Week 16 through Week 48 for assessments of efficacy, safety, and other measures of QOL/health status/work productivity. An SFU visit will be conducted 20 weeks after the last dose of IMP for participants who do not enter the extension study, or who are prematurely withdrawn from the study.

4.1.1 Screening Period (Weeks -5 to 0)

The Screening Period will last a minimum of 14 days up to a maximum of 5 weeks prior to randomization.

4.1.2 Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)

Eligible study participants will be randomized in a 2:2:2:1 ratio as noted in the Study Schema to:

- Bimekizumab 320mg Q2W from Weeks 0 to 48
- Bimekizumab 320mg Q4W from Weeks 0 to 48
- Bimekizumab 320mg Q2W to Week 16, continuing on 320mg Q4W from Weeks 16 to 48
- Placebo to Week 16, continuing on bimekizumab 320mg Q2W from Weeks 16 to 48

4.1.3 Safety Follow-up Visit

All study participants not continuing in the extension study, including those withdrawn from IMP as well as those completing all study visits, will have an SFU Visit 20 weeks after their final dose of IMP.

4.1.4 Visit Windows

Visit windows are ± 3 days (based on the date of the first dose). The minimum number of days between 2 consecutive injection visits is 8 days (eg, Visit 2 +3 days occurs on Day 17; Visit 4 -3 days occurs on Day 25). However, the minimum of 8 days between injections may be used only 1 time in the Initial Treatment Period and 1 time in the Maintenance Treatment Period, if needed. The study participant should be dosed according to the administration schedule thereafter. The 20-week SFU Visit window is ± 7 days (based on the date of the final dose).

4.2 Scientific rationale for study design

A randomized, double-blind, placebo-controlled study design has been selected to demonstrate efficacy and safety of bimekizumab for regulatory approval. The study population will include adults with moderate to severe HS. The inclusion and exclusion criteria were designed to ensure the safety of study participants, and to enroll a broad HS study participant population representative of clinical practice in terms of disease severity and morbidity (physical disability and discomfort) that warrants therapy with a systemic agent. Considering that study participants with moderate to severe HS are treated with different antibiotics, systemic tetracyclines have been selected as the most appropriate class of antibiotics for the study based on the current therapeutic guidelines for HS.

The primary efficacy outcome measure (HiSCR₅₀) is a validated clinical outcome measure for evaluating efficacy in study participants with moderate to severe HS.

In addition, a core domain set for HS study outcome established for HS calls for the concurrent measurement of 5 core outcome domains agreed by both patients and health care providers: pain, physical signs, HS specific quality of life, global assessment and progression of course. A sixth domain, symptoms, has also been recommended by the Steering Group because it received strong support from the patient stakeholder group (Thorlacius, 2018).

The Screening Period is included to ensure eligibility criteria are met, including collection of laboratory data, verification that the doses of concomitant and allowable medications are stable, and to enable washout of any medications not permitted for use during the study.

The randomization allocation and sample sizes have been selected to (1) maximize exposures to bimekizumab test doses/regimens, (2) ensure adequate power to demonstrate superiority of bimekizumab to placebo for the primary endpoint (HiSCR₅₀) at Week 16, and (3) have sufficient sample size to detect statistically significant differences between treatments as specified in the secondary endpoints at Week 16.

An initial treatment period of 16 weeks will be used to demonstrate the efficacy of bimekizumab over placebo. The 32-week Maintenance Treatment Period will collect information on safety and efficacy beyond initial treatment.

The Maintenance Treatment Period is designed to assess the durability of response of the study endpoints and to provide sufficient longer-term (up to 48 weeks, including the Initial Treatment Period) safety and exposure to bimekizumab for regulatory filings. Continuous (up to 48 weeks) exposure to both the 320mg bimekizumab Q2W and 320mg bimekizumab Q4W dose regimens allow for the following assessments:

- Durability of response and longer-term safety
- Optimal dosing interval for maintenance treatment

This period will also allow study participants who received placebo in the Initial Treatment Period to begin receiving bimekizumab at Week 16 in the randomized, controlled Maintenance Treatment Period of the study.

4.3 Justification for dose

The pathophysiology of HS is an active area of research, with investigations targeting identification of the cytokines and immune pathways in HS. Recent reviews on the subject indicate the potential diversity of these inflammatory pathways and mediators, and the impact on treatment response to various pharmacologic interventions. As concluded by Frew, Hawkes, and Krueger, no current schema accurately predicts treatment efficacy to date (Prens, 2015; Frew, 2018). Furthermore, the inflammatory burden in HS seems greater than other autoinflammatory conditions affecting the skin (Van der Zee, 2011; Riis, 2015; Martorell, 2015). These data, and the current pharmacologic treatment of HS suggest that a more intensive dosing regimen (ie, dose level and/or frequency of administration) may be needed for the treatment of HS.

Consistent with the above literature, a 320mg Q2W dose regimen of bimekizumab is being evaluated in the Phase 3 program for HS. This dose regimen is higher than those used in other

indications currently in Phase 3 development for bimekizumab; however, HS0001 results revealed that bimekizumab 320mg Q2W demonstrated consistent, clinically meaningful efficacy in the treatment of HS when compared to placebo, with safety results consistent with those of studies of bimekizumab in other indications in development. In addition, pharmacokinetic (PK) data from HS0001 demonstrated that study participants with HS have a lower exposure to bimekizumab than study participants with PSO given the same dose and regimen, thus necessitating higher bimekizumab doses for HS (maximum monthly bimekizumab dose of 640mg).

In addition to the 320mg Q2W dose used in HS0001, the highest dose being used in other bimekizumab indications (PSO dose regimen of 320mg Q4W) is being used in this study. The assessment of both the Q2W and Q4W dosing frequencies in this study will help determine the optimal monthly bimekizumab dose required to sustain efficacy with long-term (maintenance) treatment.

4.4 End of study definition

A study participant will be considered to have completed the study if he or she completed the Week 48 visit.

The end of the study is defined as the date of the last scheduled procedure shown in the Schedule of Activities ([Table 1-1](#)) for the last study participant in the study globally, including the SFU, as applicable.

5 STUDY POPULATION

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

5.1 Inclusion criteria

Participants are eligible to be included in the study only if all of the following criteria apply:

Age

1. Participant must be at least 18 years of age, at the time of signing the informed consent. If a study participant is under the local age of consent and is at least 18 years of age, written informed consent will be obtained from both the study participant and the legal representative.

Type of participant and disease characteristics

- 2a. Study participants must have a diagnosis of HS based on clinical history and physical examination for at least 6 months prior to the Baseline visit; diagnosis must be verifiable through medical notes and documentation.
3. Study participant must have HS lesions present in at least 2 distinct anatomic areas (eg, left and right axilla), 1 of which must be at least Hurley Stage II or Hurley Stage III at both the Screening and Baseline visits.
4. Study participant must have moderate to severe HS defined as a total of ≥ 5 inflammatory lesions (ie, number of abscesses plus number of inflammatory nodules) at both the Screening and Baseline visits.

5a. Study participant must have had a history of inadequate response to a course of a systemic antibiotics for treatment of HS at the Screening Visit as assessed by the Investigator through study participant interview and review of medical history; inadequate response must be verifiable through medical notes and documentation. Study participants who meet any of the following are NOT automatically excluded from the study:

- Demonstrated intolerance to (or during therapy became intolerant to) systemic antibiotics
- Had a contraindication to systemic antibiotics
- Responded to course(s) of systemic antibiotic(s) and subsequently exhibited recurrence after discontinuation of the antibiotic

Sex

6. Males and females may be study participants.

- A female study participant is eligible to participate if she is not pregnant (see Appendix 4 [Section 10.4]), not breastfeeding, and at least one of the following conditions applies:
 - Not a woman of childbearing potential (WOCBP) as defined in Appendix 4
 - OR
 - A WOCBP who agrees to follow the contraceptive guidance in Appendix 4 during the treatment period and for at least 20 weeks after the last dose of IMP.

Informed consent

7. Study participant was capable of giving signed informed consent as described in Appendix 1 which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in this protocol.

5.2 Exclusion criteria

Study participants are excluded from the study if any of the following criteria apply:

Medical conditions

1. Study participant has any medical or psychiatric condition that, in the opinion of the Investigator, could jeopardize or would compromise the study participant's ability to participate in this study as determined by the Investigator based on protocol-required assessments.

HS, Skin-Specific, and Other Inflammatory Disease

2. Study participant has a draining tunnel count of >20 at the Baseline Visit.
3. Study participant has any other active skin disease or condition (eg, bacterial cellulitis, candida intertrigo, extensive condyloma) that may, in the opinion of the Investigator, interfere with the assessment of HS.
4. Study participant has a diagnosis of sarcoidosis, systemic lupus erythematosus, or active IBD. Note: Study participants with a diagnosis of Crohn's disease or ulcerative colitis are allowed if they have no active symptomatic disease at Screening or Baseline.

5. Study participant has a primary immunosuppressive condition, including taking immunosuppressive therapy following an organ transplant, or has had a splenectomy.

Other Medical Conditions

6. Female study participant who is breastfeeding, pregnant, or plans to become pregnant during the study or within 20 weeks following the final dose of IMP.
7. Study participant has an active infection or history of infection(s) as follows:
- Any infection requiring systemic treatment within 14 days prior to Baseline
 - A serious infection, defined as requiring hospitalization or intravenous anti-infective(s) within 2 months prior to the Baseline Visit
 - A history of opportunistic, recurrent, or chronic infections that, in the opinion of the Investigator, might cause this study to be detrimental to the study participant. Opportunistic infections are infections caused by uncommon pathogens (eg, *Pneumocystis jirovecii*, cryptococcosis), or unusually severe infections caused by common pathogens (eg, cytomegalovirus, herpes zoster)
8. Study participant has any of the following:
- Known active TB disease
 - History of active TB involving any organ system unless adequately treated according to World Health Organization/Centers for Disease Control and Prevention therapeutic guidance and proven to be fully recovered upon consult with a TB specialist
 - Latent TB infection (LTBI). Participants with LTBI diagnosed during Screening must have completed a course of prophylaxis prior to IMP dosing. Participants can be rescreened after completion of a full course of prophylaxis plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline to avoid any interference with the study efficacy measurements (eg, concomitant antibiotics). Prophylaxis should be in accordance with applicable clinical guidelines and TB specialist judgment based on the origin of infection.
 - High risk of exposure to TB infection
 - Current pulmonary nontuberculous mycobacterial (NTM) infection or history of pulmonary NTM infection unless proven to be fully recovered

Note: For further information relating to definitions of known active TB, past history of TB, LTBI, high risk of acquiring TB infection and NTM infection refer to Section 8.2.6.

9. Study participant has an acute or chronic hepatitis B virus, hepatitis C virus (HCV), or human immunodeficiency virus (HIV) infection. Study participants who have evidence of, or tested positive for, hepatitis B or hepatitis C will be excluded. A positive test for hepatitis B virus is defined as: 1) positive for hepatitis B surface antigen, or 2) positive for anti-hepatitis B core antibody. A positive test for HCV is defined as: 1) positive for hepatitis C antibody, and 2) positive via a confirmatory test for HCV (for example, HCV polymerase chain reaction).

10. Study participants with concurrent malignancy are excluded. Study participants with a history of malignancy within the past 5 years prior to the Screening Visit are excluded, EXCEPT if the malignancy was a cutaneous squamous or basal cell carcinoma, or in situ cervical cancer that has been treated and is considered cured.
11. Study participant has a history of a lymphoproliferative disorder including lymphoma or current signs and symptoms suggestive of lymphoproliferative disease.
12. Study participant has had major surgery within the 3 months prior to the Baseline Visit, or has planned major surgery after entering the study.
13. Study participant has any systemic disease (ie, cardiovascular, neurological, renal, liver, metabolic, gastrointestinal, hematological, immunological, etc.) considered by the Investigator to be uncontrolled, unstable, or likely to progress to a clinically significant degree during the course of the study.
14. Study participant has had a myocardial infarction or stroke within the 6 months prior to the Screening Visit.
15. Study participant has a history of chronic alcohol or drug abuse within 6 months prior to Screening as evaluated by the Investigator based on medical history, interview, and/or results of the Screening urine drug screen.
- 16a. Study participant has the presence of active suicidal ideation, or positive suicide behavior using the “Screening” version of the electronic Columbia Suicide Severity Rating Scale (eC-SSRS) with either of the following criteria:
 - Study participant has a history of a suicide attempt within the 5 years prior to the Screening Visit. Study participants with a history of a suicide attempt more than 5 years ago should be evaluated by a mental healthcare practitioner before enrolling into the study.
 - Suicidal ideation in the past month prior to the Screening Visit as indicated by a positive response (“Yes”) to either Question 4 or Question 5 of the “Screening” version of the eC-SSRS.
17. Study participant has presence of moderately severe major depression or severe major depression indicated by a score of ≥ 15 using the screening Patient Health Questionnaire Depression Module (PHQ-9). Medication used to treat depression should be stable for 8 weeks prior to Baseline.
18. Study participant has a known hypersensitivity to any components of bimekizumab or comparative drugs as stated in this protocol.

Prior/Concomitant therapy

- 19a. Study participant has had prior treatment with an IL-17 biologic response modifier or has participated in IL-17 biologic response modifier study unless an appropriate washout has been performed since the last dose of IMP (within 6 months prior to the Baseline Visit or 5 half-lives [whichever is greater]).
20. Study participant received prescription topical therapies for the treatment of HS within 14 days prior to the Baseline Visit.

21. Study participant is currently receiving systemic nonbiologic or biologic therapies for HS with potential therapeutic impact for HS. Note: If study participant received systemic nonbiologic or biologic therapies for HS and stopped these treatments, washout periods should be applied as shown in [Table 6–3](#). Note: this does not apply to study participants who may be eligible for randomization into the antibiotic strata.
22. If study participant is using concomitant, non-opioid analgesics for HS-related or non-HS-related pain as permitted by protocol, they should be on a stable (scheduled) dose for at least 14 days prior to the Baseline Visit and anticipate continuing that dose through Week 16 unless a decrease in dose is warranted based on symptoms. Opioid analgesics are excluded. Note: As needed (PRN) use is not considered a stable dose, but (for example) taking a nonsteroidal anti-inflammatory drug (NSAID) 3 times per week, every week is considered a stable dose.
23. Study participant has received any live (including attenuated) vaccination within the 8 weeks prior to the Baseline Visit (eg, inactivated influenza and pneumococcal vaccines are allowed, but nasal influenza vaccination is not permitted). Live vaccines are not allowed during the study, including the SFU Period (20 weeks after the last dose of IMP).
24. Study participant has received Bacillus Calmette-Guerin vaccination within 1 year prior to IMP administration.

Prior/Concurrent clinical study experience

25. Study participant has previously participated in this study or study participant has previously been assigned to treatment in a study of the medication under investigation in this study, and received at least 1 dose of IMP (including placebo).
26. Study participant is currently participating in another study of a systemic medication under investigation, including SFU. Study participant must be washed out of the medication as indicated in [Table 6–3](#).
27. Study participant is currently participating in another study of a topical medication under investigation, including SFU. Study participant must be washed out of the medication for 4 weeks prior to the Baseline Visit.
28. Study participant is currently, or was within the 4 weeks prior to the Baseline Visit, participating in another study of a medical device under investigation.

Diagnostic assessments

29. Study participant has laboratory abnormalities at Screening, including any of the following:
 - $\geq 3 \times$ the upper limit of normal (ULN) of any of the following: alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP)
 - Bilirubin $> 1.5 \times \text{ULN}$ (isolated bilirubin $> 1.5 \times \text{ULN}$ is acceptable if bilirubin is fractionated and direct bilirubin $< 35\%$)
 - White blood cell count $< 3.00 \times 10^3/\mu\text{L}$
 - Absolute neutrophil count $< 1.5 \times 10^3/\mu\text{L}$

- Lymphocyte count <500 cells/ μ L
- Hemoglobin <8.5g/dL

Note: Individual screening tests for which the results are in error, borderline, or indeterminate for inclusion in the study can be repeated once for confirmation during the Screening Period. Upon retesting, study participants whose results remain outside this threshold should not be randomized.

30. Study participant has any other laboratory abnormality, which, in the opinion of the Investigator, will prevent the study participant from completing the study or will interfere with the interpretation of the study results.

Other exclusions

31. Study participant is a UCB employee or is an employee of third-party organizations involved in the study.
32. Study participant and/or his or her immediate family member is an employee, volunteer, or other worker at the investigative site either affiliated or not affiliated with this study. Immediate family is defined as a spouse, parent, child, or sibling whether biological or legally adopted.

5.3 Lifestyle restrictions

Not applicable to this study.

5.4 Screen failures

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomized to study treatment. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any serious adverse events (SAE). Individuals who do not meet the criteria for participation in this study (screen failure) may be rescreened, following discussion with the Medical Monitor or Sponsor's study physician.

Participants who are rescreened should be assigned a new participant number for rescreening.

A study participant may be rescreened 1 time for reasons including, but not limited to, the following:

- Individual laboratory screening tests for which the results are exclusionary can be retested (eg, tests that result in ALT, AST, or ALP up to 25% above the exclusion limit may be repeated once for confirmation. Test can also be repeated during rescreening.) Of note, repetition of laboratory screening tests within the Screening Period is permitted for technical reasons (eg, frozen sample, expired laboratory kit) without contacting the Medical Monitor.
- Eligibility assessments that could not be completed as planned (eg, for technical reasons) within the defined Screening Period of 5 weeks without approval by Medical Monitor.
- Abnormal ECG results.

- Did not meet the required washout period for concomitant medications.
- Study participant needs to complete a full course of antibiotic therapy for LTBI plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline as described in Exclusion Criterion 8 (Section 5.2).
- If the study participant requires an incision and drainage procedure for a HS lesion(s) during the Screening Period, the study participant should be screen failed. The participant can be rescreened when the lesion is considered healed. The study participant must have completed antibiotics/analgesic treatment if required for the procedure before rescreening as described in Table 6–3.

Study participants who fail to meet the eligibility criteria for PHQ-9, eC-SSRS, or the TB questionnaire are not allowed to be rescreened.

The Medical Monitor must be contacted for confirmation of rescreening/retesting in all other cases.

6 STUDY TREATMENTS

Study treatment is defined as any investigational treatment, marketed product, placebo, or medical device intended to be administered to a study participant according to the study protocol.

6.1 Treatments administered

Eligible study participants will be randomized in a 2:2:2:1 ratio as noted in the Study Schema (Figure 1-1).

A summary of the treatments administered is provided in Table 6–1.

Table 6–1: Study medications administered

ARM Name	Bimekizumab	Placebo
Intervention name	Bimekizumab	Placebo
Type	Biologic	Drug
Dose formulation	Solution for injection (pre-filled 1-mL syringe)	Solution for injection (pre-filled 1-mL syringe)
Unit dose strengths	160mg/mL	0.9% sodium chloride aqueous solution (physiological saline, preservative free) of US Pharmacopoeia/European Pharmacopoeia quality appropriate for injection; same volume to maintain blinding to the study participant
Dosage levels	320mg	Not applicable
Route of administration	Subcutaneous injection	Subcutaneous injection
Use	Experimental	Placebo comparator
Investigational Medicinal Product and Non-Investigational Medicinal Product	Investigational Medicinal Product	Investigational Medicinal Product
Sourcing	Provided centrally by the Sponsor.	Provided centrally by the Sponsor.
Packaging and labeling	Study Intervention will be provided in a kit. Each kit will be labeled as required per country requirement	Study Intervention will be provided in a kit. Each kit will be labeled as required per country requirement
Current/Former names or aliases	Bimekizumab	Not applicable

Because of differences in the dosing schedules and in order to maintain blinding, all study participants will receive 2 injections sc every 2 weeks from Week 0 to Week 46 as depicted in [Table 6–2](#).

Table 6–2: Dosing scheme

Week Dose Assignment	Initial Treatment Period (weeks after first dose)									Maintenance Treatment Period (weeks after first dose)															
	Baseline 0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	
Bimekizumab 320mg Q2W/Q2W	●●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	
Bimekizumab 320mg Q4W/ Q4W	●●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	
Bimekizumab 320mg Q2W/ 320mg Q4W	●●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	○ ●	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	
Placebo/bimekizumab 320mg Q2W	○○	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	

Q2W=every 2 weeks; Q4W=every 4 weeks

Notes: A bimekizumab 160mg injection is depicted by a black circle (●). A placebo injection is depicted by a white circle (○).

6.2 Preparation, handling, storage, and accountability requirements

The Investigator or designee must confirm appropriate temperature conditions have been maintained during transit for all study treatment received and any discrepancies are reported and resolved before use of the study treatment.

Only study participants enrolled in the study may receive study treatment and only authorized site staff may supply or administer study treatment. All study treatments must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the Investigator and authorized site staff.

The Investigator, institution, or the head of the medical institution (where applicable) is responsible for study treatment accountability, reconciliation, and record maintenance (ie, receipt, reconciliation, and final disposition records).

In case an out-of-range temperature is noted, it must be immediately reported as per instructions contained in the IMP Handling Manual.

Further guidance and information for the final disposition of unused study treatment are provided in the IMP Handling Manual.

6.2.1 Drug accountability

A Drug Accountability form will be used to record IMP dispensing and return information on a by-participant basis and will serve as source documentation during the course of the study. Details of any IMP lost, damaged (due to breakage or wastage), not used, partially used, disposed of at the study site, or returned to the Sponsor or designee must also be recorded on the appropriate forms. All supplies and pharmacy documentation must be made available throughout the study for UCB (or designee) to review.

The Investigator (or designee) is responsible for retaining all used, unused, and partially used containers of IMP until returned or destroyed.

The Investigator may assign some of the Investigator's duties for drug accountability at the study site to an appropriate pharmacist/designee.

The Investigator must ensure that the IMP is used only in accordance with the protocol.

Periodically, and/or after completion of the clinical phase of the study, all used (including empty containers)/partially used, unused, damaged, and/or expired IMP must be reconciled and either destroyed at the site according to local laws, regulations, and UCB standard operating procedures or returned to UCB (or designee). Investigational medicinal product intended for the study cannot be used for any other purpose than that described in this protocol.

6.3 Measures to minimize bias: randomization and blinding

6.3.1 Randomization and numbering of participants

An interactive response technology (IRT) system will be used for assigning eligible study participants to a treatment regimen (as applicable) based on a predetermined production randomization and/or packaging schedule provided by UCB (or designee). The randomization schedule will be produced by the IRT vendor. The IRT will generate individual assignments for kits of IMP, as appropriate, according to the visit schedule.

To enroll a study participant at Screening, the Investigator or designee will contact the IRT and provide brief details about the study participant to be enrolled. Each study participant will receive a unique number assigned at Screening that serves as the study participant identifier throughout the study. The study participant number will be required in all communication between the Investigator or designee and the IRT regarding a particular study participant. Study participant numbers and kit numbers will be tracked via the IRT.

To randomize a study participant, the Investigator or designee will contact the IRT and provide brief details about the study participant to be randomized. The IRT will automatically inform the Investigator or designee of the study participant's randomization number. The IRT will allocate kit numbers to the study participant based on the study participant number during the course of the study. The randomization number must be incorporated into the case report form (CRF).

6.3.2 Procedures for maintaining and breaking the treatment blind

6.3.2.1 Maintenance of study treatment blind

All study participant treatment details will be allocated and maintained by the IRT system.

The IRT provider will receive the randomization code at the start of the study.

Due to differences in presentation between bimekizumab and placebo treatments, special precautions will be taken to ensure study blinding; study sites will have blinded and unblinded personnel. Bimekizumab and placebo injections will be administered at the investigational sites by unblinded, dedicated study personnel according to the site-specific blinding plan. Unblinded study personnel will be responsible for recording the administration information on source documents, and administration of the IMP as sc injections. Study site pharmacists or other suitably qualified site personnel who are responsible for preparation and administration of IMP treatments will have access to treatment allocations for individual study participants via the IRT. The unblinded pharmacy monitors from the Contract Research Organization (CRO), and the UCB Clinical Trial Supply representative will also have access to the treatment allocations and to the drug accountability records, if applicable.

The following individuals may, as necessary, have access to the randomization code as indicated:

- Members of the Data Monitoring Committee (DMC) who participate in unblinded sessions will be given information about the IMP allocation for those study participants for whom data are provided.
- The unblinded, independent CRO staff supporting preparation of the data outputs for the DMC reviews.

The unblinded study site personnel will not be involved in the study in any way other than assuring the IMP is taken from the correct kit and prepared according to the IMP-handling manual, and administering the IMP to the study participants.

In addition, high-sensitivity C-reactive protein (hs-CRP) results will not be reported to any blinded study personnel as long as the study remains blinded.

6.3.2.2 Breaking the treatment blind in an emergency situation

In the event of an emergency, it will be possible to determine to which treatment arm and dose the study participant has been allocated by contacting the IRT. The Investigator is responsible for

breaking the treatment blind in case of emergency. All sites will be provided with details of how to contact the system for code breaking at the start of the study. The Medical Monitor and/or UCB study physician or equivalent should be consulted prior to unblinding, whenever possible.

The Clinical Project Manager will be informed immediately via the IRT when a code is broken, but will remain blinded to specific treatment information. Any unblinding of the IMP performed by the Investigator must be recorded in the source documents and on the Study Termination electronic CRF (eCRF) page.

Inadvertent unblinding will be listed as a major protocol deviation.

6.4 Treatment compliance

During the double-blind Initial Treatment and Maintenance Treatment Periods, IMP will be administered in the clinic and compliance will be recorded at the visit by study personnel in the eCRF. Drug accountability must be recorded on the Drug Accountability form.

6.5 Concomitant medications/treatments

6.5.1 Permitted concomitant treatments (medications and therapies)

The following concomitant medications are permitted during the study:

- Wound care: Concomitant use of wound care dressings on HS wounds is allowed; however, options are limited to alginates, hydrocolloids, and hydrogels and use of these will be recorded in the eCRF.
- Lesion care: Concomitant use of saline, water, and/or Vaseline (petroleum jelly) is allowed for care of skin lesions and use of these will be recorded in the eCRF.
- Analgesic therapy:
 - Study participants will be required to wash out of all analgesics for HS-related pain 14 days prior to Baseline. However, if a study participant is on a stable (scheduled) dose of a non-opioid analgesic for HS-related pain, or for a non-HS medical condition (eg, osteoarthritis, neuropathic pain), the study participant may continue the analgesic. Opioid analgesics (including tramadol) are excluded for any indication.

Notes: (1) Dose should be stable for 14 days prior to Baseline, and is anticipated to remain stable throughout study participation. (2) Dosing PRN is not considered stable, but (for example) taking an NSAID 2 or 3 times per week every week is considered a stable dose.

- If a study participant's pain (HS-related or non-HS-related) worsens after Baseline, the study participant may initiate analgesic therapy at any time and/or per local labeling as follows: For HS-related pain, permitted analgesics are limited to ibuprofen at a dose of up to 800mg orally every 6 hours, not to exceed 3.2g/24 hours; and/or acetaminophen/paracetamol as per local labeling. For non-HS-related pain, initiation of any new analgesic/treatment must not include exclusionary medications (eg, opioids and tramadol), and must be recorded on the eCRF.
 - All analgesic use (start dates, end dates, dose, reason) will be recorded on the eCRF.

- Antibiotic therapy:
 - For study participants entering the study in the antibiotic strata, they should be on a stable dose and regimen of doxycycline, minocycline, or an equivalent systemic tetracycline for 28 days prior to Baseline (Visit 2). The dose and regimen should remain stable throughout study participation, but at least through Week 16. Antibiotics taken on a PRN basis are not considered as a stable dose. After Week 16, participants may receive an antibiotic if required in the judgement of the Investigator. Also see Section 6.5.3.1 for details on systemic antibiotic rescue medication.
 - All antibiotic use (start dates, end dates, dose, reason) will be recorded on the eCRF.

6.5.2 Prohibited concomitant treatments (medications and therapies)

The following concomitant medications and therapies are prohibited during the study (also see Table 6–3):

- All biologic therapy with a potential therapeutic impact on the disease being studied, including those listed in Table 6–3.
- Phototherapy (psoralen and ultraviolet A and/or ultraviolet B) or photochemotherapy.
- Immunomodulatory therapy, including topical or systemic steroids except as noted in Section 6.5.3.1 (Rescue Medications/Lesion Intervention), and Table 6–3.
- Topical and systemic therapies for HS (see Table 6–3).
- Surgical or laser intervention for an HS lesion except as outlined in Section 6.5.3.1 (Rescue Medications/Lesion Intervention).

Table 6–3: Exclusions for prior medications

Drug class	Exclusion criteria
<u>Systemic antibiotics</u>	Used within 2 weeks prior to the Baseline Visit. Note: See exception for permitted, stable doses of antibiotics in Section 6.5.1.
<u>Systemic retinoids</u>	Used within 4 weeks prior to the Baseline Visit
<u>Systemic treatment (non-biologic)</u> <ul style="list-style-type: none"> • Apremilast • Systemic immunosuppressant agents (eg: methotrexate, cyclosporine, azathioprine, thioguanine) • Systemic fumarate • Systemic oral or injectable corticosteroids • Phototherapy and radiotherapy (eg, psoralen and ultraviolet A and/or ultraviolet B) or photoradio/chemotherapy 	Used within 4 weeks prior to the Baseline Visit. Note: See exception for permitted, stable doses of antibiotics in Section 6.5.1.
<u>Anti-tumor necrosis factors (including biosimilars)</u> adalimumab, etanercept, certolizumab, golimumab, infliximab	Used within 12 weeks prior to the Baseline Visit. Note: For etanercept, used within 1 month prior to the Baseline Visit.
<u>Other biologics</u> Abatacept Anakinra Natalizumab Belimumab Tocilizumab Efalizumab Or other biologics approved by regulatory agencies after the protocol is approved	Used within 12 weeks prior to the Baseline Visit Note: for other biologics approved by regulatory agencies after the protocol is approved: Washout of 6 months prior to the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater.
Secukinumab, brodalumab, ixekizumab and other IL-17 inhibitors approved by regulatory agencies after the protocol is approved	Washout of 6 months prior to the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater.
IL-12, IL-23 inhibitors: Ustekinumab Risankizumab Tildrakizumab Guselkumab Or other biologics approved by regulatory agencies after the protocol is approved	Used within 6 months prior to the Baseline Visit Note: for other biologics approved by regulatory agencies after the protocol is approved: Washout of 6 months prior to the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater.

Table 6–3: Exclusions for prior medications

Drug class	Exclusion criteria
<u>Janus kinase inhibitors</u> Tofacitinib Baricitinib Filgotinib Upadacitinib Or other janus kinase inhibitors approved by regulatory agencies after the protocol is approved	Used within 12 weeks of the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater
Any other systemic HS drug under investigation (or approved after the protocol is approved)	Used within 12 weeks or 5 half-lives prior to the Baseline Visit, whichever is greater
Rituximab	Used within 2 years of the Baseline Visit
Topical drugs for HS (including intralesional corticosteroids, over-the-counter and prescription drugs, as well as disinfectants for skin lesions, eg, chlorhexidine, povidone iodine)	Used within 14 days of the Baseline Visit
Topical corticosteroids (in HS-affected areas) for dermatological use	Used within 14 days of the Baseline Visit. Note: Topical steroids in non-HS affected areas are permitted.
Herbal medications for HS	Used within 14 days of the Baseline Visit
Vaccines	Administration of live (including attenuated) vaccines is not allowed within 8 weeks prior to Baseline, during the conduct of the study, and for 20 weeks after the final dose of IMP (see Exclusion Criteria #23 and #24). Administration of inactivated vaccines is allowed during the study at the discretion of the Investigator. Administration of any other vaccine not mentioned above may be allowed following discussion with the Medical Monitor.
Analgesics	See Section 6.5.1.
Spironolactone	Permitted if indicated for non-HS-related condition (eg, polycystic ovary syndrome); and if study participant meets all other entry criteria. Study participant must be on a stable dose for at least 28 days prior to the Baseline visit.

Table 6–3: Exclusions for prior medications

Drug class	Exclusion criteria
Metformin	Permitted if indicated for non-HS-related condition (eg, diabetes mellitus); and if study participant meets all other entry criteria. Study participant must be on a stable dose for at least 28 days prior to the Baseline visit.
Finasteride and other 5 α -reductase inhibitors	Permitted if indicated for non-HS-related condition (eg, benign prostatic hypertrophy); and if study participant meets all other entry criteria. Study participant must be on a stable dose for at least 28 days prior to the Baseline visit.

HS=hidradenitis suppurativa; IL=interleukin; IMP=investigational medicinal product

6.5.3 Rescue medication

The Sponsor will not supply rescue medication. The following rescue medications may be used:

6.5.3.1 Antibiotic Rescue Medication/Antibiotics

Any systemic antibiotic that is initiated (new antibiotic or change in the dose/type of current antibiotic) on or after Baseline (first day of study drug administration) will be considered rescue medication for both the Initial Treatment Period and Maintenance Treatment Period. If a newly initiated systemic antibiotic (or increase in dose/type of antibiotic) is required during the Initial Treatment Period based on disease flare or other extenuating circumstances, the Investigator should discuss the decision with the Medical Monitor.

6.5.3.2 Rescue Medications/Lesion Intervention

There are no absolute restrictions on the use of rescue medications for study participants whose HS deteriorates during the study. While the objectives of the study should be protected as much as possible through observance of the restrictions detailed above in Section 6.5, the well-being of the study participant will always take priority; study participants should be managed as deemed appropriate by the Investigator.

In the event an acutely painful lesion occurs that requires an immediate intervention, Investigators will have the option to perform interventions. Interventions can include analgesics for a limited period of time (see below), intralesional injections of triamcinolone, and/or incision and drainage of the abscess. Intralesional injections of triamcinolone (up to 20mg across all lesions at a given visit, and using a concentration of no more than 20mg/mL [suspension for injection]) must be consistent with the maximum number of interventions described below and clinical practice. Concomitant use of wound care dressings is permitted; however, options are limited to alginates, hydrocolloids, and hydrogels. Concomitant medications associated with the lesion intervention(s) must be captured in source documents and on the appropriate eCRF.

Any analgesic that is initiated (new analgesic, new class of analgesic, increased dose of an analgesic stable since Baseline, regardless of duration of treatment) after Baseline (first day of study drug administration) will be considered rescue pain medication for the Initial Treatment Period and Maintenance Treatment Period (Weeks 0 to 48).

A total of 2 protocol-allowed interventions are permissible during the Initial Treatment Period (from Baseline Visit to Week 16). Do not include analgesic rescue treatment in the number of protocol-allowed interventions. An intervention can occur on maximally 2 different lesions at the same visit, or on the same lesion at 2 different study visits. The same lesion cannot be treated 2 times at the same visit. If a study participant requires more than 2 interventions within the first 16 weeks of the study, then the study participant should be discontinued from the study.

During the Maintenance Treatment Period (Weeks 16 to 48), a maximum of 2 interventions every 4 weeks are permitted. Do not include analgesic rescue treatment in the number of protocol-allowed interventions. An intervention can occur on 2 different lesions at the same visit or on the same lesion at 2 different study visits. Within each 4-week period, the same type of intervention cannot be used 2 times on the same lesion. If a study participant requires more than 2 interventions within a 4-week period, or has 2 of the same interventions on the same lesion within that period, then the study participant should be discontinued from the study.

6.6 Dose modification

Dose modification is not applicable in this study.

6.7 Criteria for study hold or dosing stoppage

UCB reserves the right to temporarily suspend or prematurely discontinue this study either at a single site, multiple sites, or at all sites at any time for reasons including, but not limited to, safety or ethical issues, inaccurate or incomplete data recording, noncompliance, or unsatisfactory enrollment with respect to quality or quantity. [REDACTED]

If the study is prematurely terminated or suspended, UCB (or its representative) will inform the Investigators/institutions and the regulatory authority(ies) of the termination or suspension and the reason(s) for the termination or suspension, in accordance with applicable regulatory requirement(s). The Institutional Review Board (IRB)/ Independent Ethics Committee (IEC) should also be informed and provided with reason(s) for the termination or suspension by the Sponsor or by the Investigator/institution, as specified by the applicable regulatory requirement(s). In addition, arrangements will be made for the return or destruction of all unused IMP and other material in accordance with UCB procedures for the study.

Detailed procedures for reporting SAEs and other safety events which may meet study hold criteria are provided in Appendix 8.

6.8 Treatment after the end of the study

Study participants who complete HS0003 will have the option of enrolling in a Phase 3, multicenter, extension study (HS0005).

Study participants who elect not to enroll in HS0005 at Week 48 will be scheduled to have the SFU Visit 20 weeks after the final injection of IMP. During the SFU, if study participants' HS deteriorates, the Investigator may consider standard of care for HS treatment after discussion with the Medical Monitor or UCB study physician. Note that the half-life of bimekizumab must be considered in selection of appropriate HS treatments during the SFU period. All concomitant

medications and HS interventions administered during the SFU will be recorded on the appropriate eCRF pages.

7 DISCONTINUATION OF STUDY MEDICATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL

7.1 Discontinuation of study medication

Study medication will be stopped if the study participant develops a medical condition, adverse event (AE), or laboratory abnormality that, in the opinion of the Investigator, compromises the safety of the study participant or his or her ability to continue participation in the study. Study participants who are discontinued from IMP should be encouraged by the Investigator to return for all scheduled visits through Week 48, and the SFU Visit (if the Week 48 Visit is ≥ 20 weeks after the final dose of IMP, the SFU Visit is not required). Any study participant who discontinues IMP but continues in the study should be discussed with the Medical Monitor or UCB study physician.

In all cases the study participant should be followed until the condition has resolved as agreed by the Investigator and the UCB study physician. Investigators should contact the Medical Monitor and/or UCB study physician, in advance whenever possible, to discuss the withdrawal of a study participant.

Study medication will be stopped if the study participant has a confirmed positive coronavirus disease 2019 (COVID-19) test result or a suspected COVID-19 infection. Study medication can be resumed after the participant's recovery from COVID-19, based on the Investigator's clinical judgement. All such cases must be discussed with the Medical Monitor or UCB study physician.

7.1.1 Study participant does not achieve partial response

If a study participant does not achieve a partial response (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at any visit from Week 32 to 46, the Investigator should contact the Medical Monitor to discuss whether the study participant should continue on study.

7.1.2 Potential drug-induced liver injury IMP discontinuation criteria

Study participants with potential drug-induced liver injury (PDILI) must be assessed to determine if IMP must be discontinued. In addition, all concomitant medications and herbal supplements that are not medically necessary should also be discontinued.

7.1.2.1 PDILI Discontinuation Criteria

The PDILI criteria below require immediate discontinuation of IMP for study participants with either of the following (see Section 10.6.2.1):

- ALT or AST $\geq 8 \times \text{ULN}$
- ALT or AST $\geq 3 \times \text{ULN}$ and coexisting total bilirubin $\geq 2 \times \text{ULN}$

Similarly, the PDILI criterion below requires immediate discontinuation of IMP for:

- Study participants with ALT or AST $\geq 3 \times \text{ULN}$ who exhibit temporally associated symptoms of hepatitis or hypersensitivity. Hepatitis symptoms include fatigue, nausea, vomiting, right

upper quadrant pain or tenderness. Hypersensitivity symptoms include fever (without clear alternative cause), rash, or eosinophilia (ie, >5%).

If a nondrug-related cause for the symptoms can be confirmed, these study participants may resume IMP administration after discussion with the Medical Monitor and/or UCB study physician, but only when the requirements for rechallenge with IMP as provided in Section 10.6.2.1 are followed.

The PDILI criterion below allows for study participants to continue on IMP at the discretion of the Investigator.

- Study participants with ALT or AST $\geq 5 \times \text{ULN}$ and $< 8 \times \text{ULN}$, total bilirubin $< 2 \times \text{ULN}$, and no eosinophilia (ie, $\leq 5\%$), with no fever, rash, or symptoms of hepatitis (eg, fatigue, nausea, vomiting, right upper quadrant pain or tenderness).

Evaluation of PDILI must be initiated as described in Section 10.6 (Appendix 6) with repeat tests performed in 2 weeks. Upon retest, if ALT or AST values have reduced to $< 5 \times \text{ULN}$, the study participant can continue with the study. However, if ALT or AST remains $\geq 5 \times \text{ULN}$ and $< 8 \times \text{ULN}$ after retest, IMP should be temporarily withheld and study participant should undergo a repeat test in 2 weeks. If ALT or AST values remain $\geq 5 \times \text{ULN}$ even after the second retest, then the study participant should be permanently withdrawn from IMP and should be followed for possible PDILI.

If study participants are unable to comply with the applicable monitoring schedule, IMP must be discontinued immediately.

Investigators should attempt to obtain information on study participants in the case of IMP discontinuation to complete the final evaluation. Study participants with PDILI should not be withdrawn from the study until investigation and monitoring are complete. All results of these evaluations and observations, as well as the reason(s) for IMP discontinuation and study participant withdrawal (if applicable), must be recorded in the source documents. The eCRF must document the primary reason for IMP discontinuation.

7.1.2.1.1 Follow-up evaluation

Potential drug-induced liver injury events require follow-up monitoring as described in Table 10–1. Monitoring should continue until liver chemistry values normalize, stabilize, or return to Baseline. Determination of stabilization is at the discretion of the Investigator in consultation with the hepatologist (as applicable) and UCB study physician, as needed.

7.1.3 Treatment interruptions

If a study participant is found to be persistently noncompliant (for example, missing 2 or more of the doses in the Initial Treatment Period or 3 or more doses during Maintenance Treatment Period) the Sponsor, in conjunction with the Investigator, will make a decision as to whether the study participant should be withdrawn from the study.

Note: Doses of IMP that were missed due to a reasonable interfering AE that does not allow administration of IMP due to safety reasons will not be considered for the evaluation of study participant discontinuation. Evaluation of the reasonability of the AE should be discussed immediately with the Medical Monitor.

Any participant who develops a clinically important infection or recurrent infections not responsive to standard therapy during the study must discontinue IMP until resolution of the infection. The Investigator should use clinical judgement in deciding whether the participant should restart IMP and contact the Medical Monitor and UCB study physician to confirm the participant's suitability for continued participation in the study.

7.2 Participant discontinuation/withdrawal from the study

Note: For female study participants, please see Section 8.3.5 for pregnancy that occurs during the study as evidenced by a positive pregnancy test.

Study participants are free to withdraw from the study at any time, without prejudice to their continued care.

A study participant may withdraw from the study at any time at his or her own request, or may be withdrawn at any time at the discretion of the Investigator for safety, behavioral, compliance, or administrative reasons.

If the study participant withdraws consent for disclosure of future information, the Sponsor may retain and continue to use any data collected before such a withdrawal of consent.

If a study participant withdraws from the study, he or she may request destruction of any samples taken and not tested, and the Investigator must document this in the site study records.

See the Schedule of Activities (Table 1–1) for data to be collected at the time of study discontinuation and follow-up and for any further evaluations that need to be completed.

If a study participant does not achieve a partial response (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at any visit from Week 32 to 46, the Investigator should contact the Medical Monitor to discuss whether the study participant should continue on study.

Study participants will be withdrawn from the study, after being encouraged to complete the Premature End of Treatment (PEOT) and the SFU Visit if either of the following events occur:

1. Study participant withdraws his or her consent.
2. The Sponsor or a regulatory agency requests withdrawal of the study participant.

Study participants should be withdrawn from IMP and encouraged by the Investigator to return for all scheduled visits through Week 48, and the SFU Visit (if the Week 48 Visit is ≥ 20 weeks after the final dose of IMP, the SFU Visit is not required) if any of the following events occur:

1. Study participant develops an illness that would interfere with his or her continued participation.
2. Study participant is noncompliant with the study procedures or medications in the opinion of the Investigator.
3. Study participant takes prohibited concomitant medications as defined in this protocol (Section 6.5.2) that may present a risk to the safety of the participant or the integrity of the study data, in the opinion of the Investigator and/or the Medical Monitor and/or UCB study physician.

4. Study participant requires more than the number of protocol-allowed lesion interventions (see Section 6.5.3.2).
5. Study participant has a clinical laboratory value meeting any of the following criteria:
 - a. Hepatotoxicity as described in Section 7.1.2.
 - b. A laboratory value meeting any of the following criteria:
 - Absolute neutrophil count $<1.0 \times 10^3/\mu\text{L}$
 - Absolute lymphocyte count $<200 \text{ cells}/\mu\text{L}$

Study participants may remain on IMP if the result is transient. A retest is required within 1 to 2 weeks at a scheduled or unscheduled visit. If the repeat absolute neutrophil count or absolute lymphocyte count is still below the allowable values, the participant must be discontinued from the IMP. If the repeat absolute neutrophil count or absolute lymphocyte count is above the allowable values, the participant may continue to receive IMP.

6. There is confirmation of a pregnancy during the study, as evidenced by a positive pregnancy test (see Section 8.3.5 for more information regarding pregnancies).
7. A study participant considered as having either a suspected new LTBI or who develops active TB or an NTM infection during the study (including but not limited to, conversion demonstrated by interferon gamma release assay [IGRA] or other diagnostic means) must be immediately discontinued from IMP.
 - The study participant must immediately be permanently withdrawn from the study if further examinations result in a diagnosis of the following:
 - active TB or
 - an NTM infection, or
 - latent TB infection and study participant does not initiate TB prophylactic therapy, prematurely discontinues TB prophylactic therapy, or, in the opinion of the Investigator or Sponsor, is noncompliant with TB prophylactic therapy.

The PEOT Visit must be scheduled as soon as possible, but not later than the next regular visit.

- If the study participant is diagnosed with LTBI during the study and desires to continue in the study, he or she must immediately discontinue IMP and start TB prophylactic therapy. After at least 4 weeks of TB prophylaxis, the IMP can be restarted after discussion with the UCB study physician regarding results of laboratory assessments, physical examination, and TB questionnaire. The full course of TB prophylaxis treatment will be completed during the study.

Confirmed active TB is an SAE and must be captured on an SAE Report Form and provided to the Sponsor in accordance with SAE reporting requirements. As with all SAEs, periodic follow-up reports should be completed as per protocol requirements until such time as the TB infection resolves.

Additional information on TB policies is provided in Section 8.2.6.

8. Study participants with newly diagnosed IBD or with IBD flares during the study must:

- Be referred, as appropriate, to a healthcare professional treating IBD, such as a gastroenterologist
- Discontinue IMP and be followed-up until resolution of active IBD symptoms

If IBD flares increase in severity or frequency during the study, the Investigator should use clinical judgement in deciding whether the study participant should continue on IMP and contact the Medical Monitor and UCB study physician to confirm the study participant's suitability for continued participation in the study.

9. Study participants must be referred immediately to a mental healthcare professional and may be withdrawn from the study based upon the Investigator's judgment of benefit/risk for:

- Active suicidal ideation as indicated by a positive response ("Yes") to Question 4 of the "Since Last Visit" version of the eC-SSRS
- Moderately severe major depression as indicated by a PHQ-9 score of 15 to 19 if this represents an increase of at least 3 points compared to the last visit

The mental health consultation must be recorded in the study participant's source documentation.

10. Study participants must be referred immediately to a mental healthcare professional and must be withdrawn from the study in case of:

- Active suicidal ideation as indicated by a positive response ("Yes") to Question 5 of the "Since Last Visit" version of the eC-SSRS.
- Any suicidal behavior since last visit.
- Severe major depression as indicated by a PHQ-9 score ≥ 20 .

The mental health consultation must be recorded in the study participant's source documentation.

Investigators should contact the Medical Monitor in advance, whenever possible, to discuss the withdrawal of a study participant from IMP or from the study.

Study participants withdrawing from the study who are not continuing for all scheduled visits through Week 48, will undergo the PEOT Visit and the SFU Visit 20 weeks after their final dose of IMP, as applicable.

The eCRF must document the primary reason for withdrawal or discontinuation.

Investigators should contact the Medical Monitor and/or UCB study physician, whenever possible, to discuss the withdrawal of a study participant in advance.

Withdrawn participants will not be replaced.

7.3 Lost to follow up

A study participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site.

The following actions must be taken if a study participant fails to return to the clinic for a required study visit:

- The site must attempt to contact the study participant and reschedule the missed visit as soon as possible and counsel the study participant on the importance of maintaining the assigned visit schedule and ascertain whether or not the study participant wishes to and/or should continue in the study.
- Before a study participant is deemed lost to follow up, the Investigator or designee must make every effort to regain contact with the study participant (at least 1 phone call and 1 written message to the participant), and document his or her effort (date and summary of the phone call and copy of the written message in the source documents), to complete the final evaluation (PEOT and SFU, as applicable). All results of these evaluations and observations, together with a narrative description of the reason(s) for removing the study participant, must be recorded in the source documents. The CRF must document the primary reason for withdrawal.

Should the study participant continue to be unreachable, he or she will be considered to have withdrawn from the study with a primary reason of lost to follow-up documented in the eCRF.

8 STUDY ASSESSMENTS AND PROCEDURES

Study procedures and their timing are summarized in the Schedule of Activities ([Table 1-1](#)).

Some study-specific assessments may be conducted remotely according to the study protocol during a pandemic or other exceptional circumstance (eg, hurricanes) due to the need to implement safety measures and guidance from regulatory authorities. In such a situation, which may be accompanied by local or global containment or other measures, sites may need to prohibit access to study participants and study-related personnel. Study participants' visits to the study site may be replaced by contingency measures. These measures are primarily established to ensure the safety of study participants during the course of the study and to maintain the study participants' treatment schedules, if the Investigator considers it appropriate. These measures include, but are not limited to, virtual visits or home-nursing visits replacing site visits, eg, telemedicine contacts or home-nursing visits when treatment and/or blood sampling is scheduled. Any study specific assessments not conducted due to such circumstances must be recorded appropriately in the source documents and eCRF. If it is related to the COVID-19 pandemic, then it must be captured in the COVID-19 impact eCRF page. The contingency measures are described in a contingency plan that will be maintained by UCB for the respective study. The contingency measures are shared with the Investigator and the respective study-related personnel as soon as there are indications that it is necessary to implement any of the measures.

Protocol waivers or exemptions are not allowed.

Immediate safety concerns should be discussed with the Medical Monitor and/or UCB study physician immediately upon occurrence or awareness to determine if the study participant should continue or discontinue study treatment.

Adherence to the study design requirements, including those specified in the Schedule of Activities, is essential and required for study conduct.

All screening evaluations must be completed and reviewed to confirm that potential study participants meet all eligibility criteria. The Investigator will maintain a screening log to record details of all study participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

Procedures conducted as part of the study participant's routine clinical management (eg, blood count) and obtained before signing of the ICF may not be utilized for screening or Baseline purposes.

The maximum amount of blood collected from each study participant over the duration of the study, including any extra assessments that may be required, will not exceed the usual volume of blood taken for a blood donation. Repeat or unscheduled samples may be taken for safety reasons or for technical issues with the samples.

8.1 Efficacy assessments

The timing for all assessments described below is specified in [Table 1–1](#).

8.1.1 Lesion count

The lesion count is defined as an assessment of all the various skin “appearances” that are termed “lesions” in HS study participants. The lesion count will include the following:

- Abscesses (circumscribed collection of purulent exudate frequently associated with swelling and other signs of inflammation, such as fluctuance, tenderness, and pain)
- Draining tunnels (fistulas/sinus tracts) (pathologic passageway connecting to the skin surface from dermis or sc tissue/pathologic passageway that develops into a channel to the skin surface that drains serous or purulent fluid, either spontaneously or by gentle palpation)
- Non-draining tunnels (fistulas/sinus tracts) (pathologic passageway connecting to the skin surface from dermis or sc tissue/pathologic passageway that develops into a channel to the skin surface that does not drain serous or purulent fluid)
- Noninflammatory nodules (nontender or minimally tender, nonerythematous nodules)
- Inflammatory nodules (a tender, erythematous, well-defined nodule. The lesion has no evidence of fluctuance. A pyogenic granuloma lesion is considered an inflammatory nodule; a papule or pustule is not considered an inflammatory nodule)
- Scars of HS lesions (enlargement or overgrowth of a scar so that it extends above the surrounding skin surface)

The data collected from the lesion count will be used for the derivation of study variables including, but not limited to HiSCR₂₅ (a 25% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count), HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀ (a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count), HS Physician's Global Assessment, AN count, and International HS Severity score system (IHS4).

8.1.1.1 Hidradenitis Suppurativa Clinical Response (HiSCR)

The HiSCR; defined as at least a 50% reduction from Baseline in the total AN count, with no increase from Baseline in abscess or draining tunnel count; was developed to address issues with

available HS scoring systems. It is a validated endpoint that is responsive to improvement in disease activity, simplifies the scoring process, and increases the sensitivity to detect HS-specific lesions (Kimball, 2014; Kimball, 2016b). HiSCR has been labeled HiSCR₅₀ in this protocol.

The HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ also will be evaluated in this study. These measures of clinical response differ from HiSCR₅₀ only in the percent decrease in AN count from Baseline.

The HiSCR_{xx} is derived by the statistical programming group based on Investigator documentation of lesion count, and does not require calculation on the part of the Investigator from the lesion count.

8.1.1.2 Hidradenitis Suppurativa Physician's Global Assessment

The HS Physician's Global Assessment is a validated 6-point scale that is used to measure improvement in inflammatory nodules, abscesses, and draining tunnels (Kimball, 2012; Zouboulis, 2015). The HS Physician's Global Assessment scale is defined by the following:

- Clear: No inflammatory or noninflammatory nodules
- Minimal: Only the presence of noninflammatory nodules
- Mild: ≥ 1 and ≤ 4 inflammatory nodules or 1 abscess or draining tunnel and no inflammatory nodules
- Moderate: ≥ 5 inflammatory nodules or 1 abscess or draining tunnel and 1 or more inflammatory nodules or 2 to 5 abscesses or draining tunnels and ≤ 10 inflammatory nodules
- Severe: 2 to 5 abscesses or draining tunnels and > 10 inflammatory nodules
- Very severe: > 5 abscesses or draining tunnels

This assessment (clear, minimal, mild, moderate, severe, or very severe) is derived based on totals across all affected body regions by the statistical programming group and does not require calculation on the part of the Investigator.

8.1.1.3 International Hidradenitis Suppurativa Severity Score System (IHS4)

The IHS4 is a validated tool to dynamically assess HS severity to be used both in real-life and the clinical trials setting (Zouboulis, 2017). The IHS4 achieved consensus among European HS Foundation members. This IHS4 score is calculated as follows: (number of nodules multiplied by 1) + (number of abscesses multiplied by 2) + [number of draining tunnels (fistulae/sinuses) multiplied by 4]. A score of 3 or less signifies mild HS, a score of 4–10 signifies moderate HS and a score of 11 or higher signifies severe HS.

The determination of IHS4 requires counting the nodules, abscesses and draining tunnels/sinus tracts.

The IHS4 score is derived by the statistical programming group, and does not require calculation on the part of the Investigator.

8.1.2 Partial response

A partial response is defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at a particular timepoint.

The partial response is derived by the statistical programming group, and does not require calculation on the part of the Investigator.

8.1.3 High-sensitivity C-reactive protein (hs-CRP)

Blood will be collected for measurement of hs-CRP. The hs-CRP data will not be sent to any blinded study personnel to protect the blinded nature of the treatment assignments and response.

8.1.4 Patient-reported outcomes

The patient-reported outcome (PRO) instruments should be completed by the study participants themselves in a quiet place. The PRO instruments to be completed at the study site, should be completed prior to all other protocol-specified assessments at each visit (including dosing on dosing days).

8.1.4.1 HS symptom measures of skin pain, smell or odor, drainage or oozing from HS lesions, and itch

8.1.4.1.1 HS symptom daily diary (HSSDD)

The 5 items on the HS Symptom Daily Diary (HSSDD) assesses patients' perception of the core symptoms of HS experienced in the past 24 hours: pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point numeric rating scale (NRS). Two items assess skin pain (ie, worst skin pain and average skin pain). The remaining 3 items assess smell or odor, itch at its worst, and amount of drainage or oozing from HS lesions.

The HSSDD will be completed daily by the study participant, at the end of the day on an electronic hand-held device from the start of Screening through the Week 16 visit.

8.1.4.1.2 HS symptom questionnaire (HSSQ)

The 4 items on the HS Symptom Questionnaire (HSSQ) assesses patients' perception of the core symptoms of HS experienced in the past 7 days - pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point NRS. The HSSQ will be completed on an electronic device during study visits (ie, Baseline and Weeks 16-48/PEOT).

8.1.4.2 Patient Global Impression of HS Severity (PGI-S-HS) and Change in HS Severity (PGI-C-HS)

The Patient Global Impression of HS Severity (PGI-S-HS) is a single item to assess study participants' perceptions of the overall severity of HS over the past 7 days (none, mild, moderate, severe, very severe). The Patient Global Impression of Change in HS Severity (PGI-C-HS) is a single item to assess study participants' perception of the change in HS since they started taking the study medication (much better, a little better, no change, a little worse, much worse). Data collected using the PGI-S-HS and PGI-C-HS will be used as anchors for interpreting change scores on the Hidradenitis Suppurativa Quality of Life (HiSQOL).

8.1.4.3 Patient Global Impression of Severity of Skin Pain (PGI-S-SP) and Change in Severity of Skin Pain (PGI-C-SP)

The Patient Global Impression of Severity of Skin Pain (PGI-S-SP) is a single item to assess study participants' perceptions of the severity of their skin pain from their HS lesions, over the past 7 days (none, mild, moderate, severe, very severe). The Patient Global Impression of Change in Severity of Skin Pain (PGI-C-SP) is a single item to assess study participants'

perceptions of change in their skin pain from their HS lesions, since they started taking the study medication (much better, a little better, no change, a little worse, much worse). PGI-S-SP and PGI-C-SP will be used to evaluate outcomes related to Skin Pain.

8.1.4.4 Hidradenitis Suppurativa Quality of Life Questionnaire (HiSQOL)

The 17 item HiSQOL questionnaire has a recall period of 7 days. The HiSQOL includes 3 subscales: symptom status, psychosocial impact, and impact on physical activities.

8.1.4.5 Dermatology Life Quality Index (DLQI)

The DLQI is a questionnaire designed for use in adult participants with inflammatory skin diseases and has been used in patients with HS (Finlay, 1998; Esmann, 2010; Basra, 2012). The DLQI is a skin disease-specific questionnaire aimed at the evaluation of how symptoms and treatment affect participants' health-related QOL. This instrument asks participants about symptoms and feelings, daily activities, leisure, work and school, personal relationships and treatment. The DLQI total score ranges from 0 to 30 with higher scores indicating lower health related QOL. In other dermatological/skin conditions, a 4-point change in the DLQI total score (DLQI response) has been reported to be meaningful for the participant (within participant minimal important difference); while a DLQI total absolute score of 0 or 1 indicates no or small impact of the disease on health related QOL.

8.1.4.6 Euro-Quality of Life 5-Dimensions, 3 levels

The EQ-5D-3L comprises 5 questions on mobility, self-care, pain, usual activities, and psychological status with 3 possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). In addition, the questionnaire includes a visual analogue scale to indicate the general health status, with 100 indicating the best health status.

8.1.4.7 Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem

The Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem (WPAI-SHP) V2.0 is a patient-reported questionnaire that assesses study participant's employment status, work absenteeism, work impairment while working (presenteeism), overall work, and daily activity impairment attributable to a specific health problem (Reilly, 1993). It has been used in several clinical studies of biologic therapy in participants with plaque PSO (Kimball, 2012; Vender, 2012).

Five out of 6 items of the WPAI-SHP are regrouped into the 4 dimensions, with scores expressed as percentage, where higher numbers indicate greater impairment and less productivity, ie, worse outcomes, as described in the WPAI-SHP scoring rules.

8.1.4.8 Treatment Satisfaction Questionnaire for Medication

The Treatment Satisfaction Questionnaire for Medication (TSQM-9) is an abbreviated 9-item version of the TSQM, excluding the side effects of medication domain. The domains included in the TSQM-9 include effectiveness (3 items), convenience (3 items) and global satisfaction (3 items). The TSQM-9 domain scores range from 0 to 100 with higher scores representing higher satisfaction (Bharmal, 2009).

8.1.5 Hurley Stage

The Hurley Stage is a severity classification for HS that was developed in 1989 and is widely used for the determination of the severity of HS (Hurley, 1989).

The Hurley Stage is defined by the following criteria:

- Stage I: Abscess formation, single or multiple, without sinus tracts and cicatrization (scarring).
- Stage II: Recurrent abscesses with tract formation and cicatrization, single or multiple, widely separated lesions.
- Stage III: Diffuse or near-diffuse involvement, or multiple interconnected tracts and abscesses across the entire area.

Hurley Stage is assigned to a given anatomic region. The overall worst Hurley Stage (ie, the highest Hurley Stage across all anatomic regions) for a given study participant at a given visit is then the study participant-level Hurley Stage. This study participant-level Hurley Stage is important for baseline stratification and disease severity assessment.

Hurley stage is included as a stratification factor for randomization.

8.2 Safety assessments

Planned timepoints for all safety assessments are provided in the Schedule of Activities (Table 1–1).

8.2.1 Physical examination

A complete physical examination will include, at a minimum, general appearance; ear, nose, and throat; eyes, hair, and skin; and assessments of the cardiovascular, respiratory, gastrointestinal musculoskeletal, and hepatic, neurological (including limb reflexes) systems, and mental status. Each physical examination also includes evaluation of signs and symptoms of active TB and risk for exposure to TB (see Section 8.2.6).

Height and weight will also be measured and recorded. The same scale for measuring body weight should be utilized throughout the study where possible.

Investigators should pay special attention to clinical signs related to previous serious illnesses.

Clinically relevant findings or worsening of previous findings since the physical examination at the Screening Visit will be recorded as AEs.

8.2.2 Vital signs

Vital signs will be measured in a sitting position after 5 minutes rest and will include body temperature (oral, axillary, otic or noncontact forehead), systolic and diastolic blood pressure, and pulse. Vital signs are to be measured prior to blood sampling, and prior to dosing, where applicable.

Vital signs will consist of single pulse and blood pressure measurements.

8.2.3 Electrocardiograms

A single 12-lead ECG will be obtained using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QT corrected for heart rate intervals.

All ECG recordings should be taken with the study participant resting in the supine position for at least 10 minutes before the recording and prior to taking blood samples or dosing.

ECG machines will be provided to study centers, and ECGs will be read by a central ECG laboratory. Full details of ECG recordings will be provided in the ECG Manual.

8.2.4 Clinical safety laboratory assessments

See Appendix 2 (Section 10.2) for the list of clinical laboratory tests to be performed.

The Investigator must review the laboratory report, document this review, and record any clinically relevant changes occurring during the study in the AE section of the CRF. The laboratory reports must be filed with the source documents. Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the Investigator to be more severe than expected for the study participant's condition.

All laboratory tests with values considered clinically significantly abnormal during participation in the study or within 20 weeks after the last dose of IMP should be repeated until the values return to normal or Baseline or are no longer considered clinically significant by the Investigator or Medical Monitor.

If such values do not return to normal/Baseline within a period of time judged reasonable by the Investigator, the etiology should be identified and the Sponsor notified.

All protocol-required laboratory assessments, as defined in Appendix 2, must be conducted in accordance with the laboratory manual and the Schedule of Activities.

If laboratory values from non-protocol specified laboratory assessments performed at the institution's local laboratory require a change in participant management or are considered clinically significant by the Investigator (eg, SAE or AE or dose modification), then the results must be recorded in the CRF.

8.2.5 Depression and suicidal risk monitoring

8.2.5.1 PHQ-9

The PHQ-9 is a multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression. The PHQ-9 scores for depression range from 0 to 27 with higher scores indicating worse state. A score of 5 to 9 is considered to be minimal symptoms of depression. A score of 10 to 14 is considered minor depression, dysthymia, or mild major depression. A score of 15 to 19 is considered to indicate moderately severe major depression, and a score ≥ 20 is considered to be severe major depression.

Refer to Section 7.2 for PHQ-9-related withdrawal criteria.

8.2.5.2 eC-SSRS

The eC-SSRS is a standardized and validated instrument developed for the assessment of the severity and frequency of suicidal ideation and behavior (Mundt, 2010; Posner, 2011). Study participants respond to standardized clinical questions that are presented in a uniform fashion.

The eC-SSRS defines 5 subtypes of suicidal ideation and behavior in addition to self-injurious behavior with no suicidal intent. The eC-SSRS takes approximately 3 to 10 minutes to complete.

Refer to Section 7.2 for eC-SSRS-related withdrawal criteria.

8.2.6 Assessment and management of tuberculosis and tuberculosis risk factors

All participants will be assessed for TB through physical examination for signs and symptoms of TB, laboratory testing (Section 8.2.4), chest x-ray (CXR) (Section 8.2.6.3.2), and TB questionnaire (Section 8.2.6.3.3).

8.2.6.1 Assessments at Screening

At Screening, all participants will have an IGRA test (QuantiFERON Gold Plus TB test is recommended), a CXR (unless already performed within 2 months of Screening; a computed axial tomography (CAT) scan of the chest at Screening or within 2 months prior to Screening is acceptable, if available), and examination for signs and symptoms of TB. In addition, the Investigator or designee will complete a TB questionnaire directed at the participants potential exposure to TB and symptoms of TB.

Study participants diagnosed with active TB during Screening will be excluded from the study.

Study participants with LTBI diagnosed during Screening must have completed a full course of prophylaxis prior to IMP dosing and can be rescreened after completion of a full course of prophylaxis plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline. (See also Section 8.2.6.3.5.)

8.2.6.2 Definitions

Study participants with known active TB disease, at high risk of acquiring TB infection, or with untreated LTBI (ie, pending anti-TB prophylactic course) or current or history of NTM infection are excluded from the study.

a. Known TB infection whether present or past is defined as:

- Active TB disease or clinical signs and symptoms strongly suggestive of TB (pulmonary or extra pulmonary).
- History of active TB disease involving any organ system or findings in other organ systems consistent with TB, unless adequately treated and proven to be fully recovered upon consult with a TB specialist.
- Any evidence by radiography or other imaging modalities consistent with previously active TB disease that is not reported in the study participant's medical history.

b. High risk of acquiring TB infection is defined as:

- Known close exposure (eg, sleeping in the same room) to another person with active TB infection within 3 months prior to Screening.
- Time spent within 3 months prior to screening in a healthcare delivery setting or institution where individuals infected with TB are housed or where the risk of transmission of infection is high.

- c. Latent TB infection is defined as an infection by *Mycobacterium tuberculosis* with:
- A positive IGRA (or 2 indeterminate IGRAs) AND
 - Chest imaging (or other imaging) negative for TB infection, AND
 - Absence of signs, symptoms (eg, evidence of organ-specific involvement), or physical findings suggestive of TB infection.
- d. Pulmonary NTM infection is defined as a group of lung or extrapulmonary infections caused by mycobacteria different from *M. tuberculosis* infections.

8.2.6.3 Assessment and reporting of TB and TB risk factors during the study

8.2.6.3.1 Physical examination

The Investigator should consider all potential sites of infection when assessing for TB during the physical examination, and other evaluations, and based on the study participant's medical or social history.

The most common primary focus of TB is the lung. Other sites may include gastrointestinal system, bone/joints, lymph glands and meninges, etc. However, in immune-compromised patients, study participants, and/or patients treated with biologics, especially tumor necrosis factors inhibitors, extra-pulmonary manifestations of TB is common compared to normal population.

Some common symptoms that the study participant may present are dependent on the primary focus of infection and may include cough, blood in sputum, night sweats, lymphadenitis, joint pain/swelling, spinal deformity, headache/confusion, abdominal pain (mimicking inflammatory bowel disease), etc. Unusual presentations should always be considered.

8.2.6.3.2 Chest x-ray for tuberculosis

Chest radiographic imaging is performed at screening and results must be available at baseline before first drug administration unless a CXR or CAT scan is available from 2 months prior to screening.

Additional CXR or other imaging test should be performed when positive signs and symptoms indicate pulmonary infection, including potential TB infection, or when close exposure to persons with TB is documented.

8.2.6.3.3 Tuberculosis questionnaire

A questionnaire entitled "Evaluation of Signs and Symptoms of Tuberculosis" has been developed by UCB (document mod-000582) to help in identifying TB risk factors in study participants; it is administered by the Investigator or their designee. For the purpose of case reporting, this questionnaire also ensures appropriate follow-up with reporters when a case of either latent TB or active TB is diagnosed. Moreover, it ensures proactive and appropriate follow-up with Investigators and study participants on treatment course.

8.2.6.3.4 IGRA Test Conversion

The IGRA is a whole-blood testing methodology for diagnosing *M. tuberculosis* infection. It has become the gold standard, but does not help in differentiating LTBI from active tuberculosis disease.

Tuberculosis test conversion is defined as a positive or indeterminate (and confirmed indeterminate on repeat) IGRA result for the current test when previous IGRA test results were negative. All study participants with positive or indeterminate IGRA test results must immediately stop IMP administration. In case of a IGRA test conversion, the study participant must be considered as having either a suspected new latent or an active TB infection and be promptly referred to an appropriate specialist (eg, pulmonologist, infectious disease specialist) for further evaluation. Additional assessments (eg, blood tests or IGRA, CXRs, or other imaging) should be performed where medically relevant and documented. Such conversions should be reported as AEs as described in the protocol. The AE term would need to be updated with final diagnosis once available.

8.2.6.3.5 Latent TB

In case the evaluation by the appropriate specialist diagnoses a new LTBI, a TB prophylactic therapy in accordance with applicable clinical guidelines should be immediately initiated.

Study participants who initiate treatment for LTBI during the Screening period must repeat initial screening laboratory parameters, all physical examinations, and questionnaires prior to randomization in the study, and must continue the full course of TB prophylactic therapy. Participants can be rescreened after completion of a full course of prophylaxis. Eligible study participants can be included after a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline.

Study participants who initiate treatment for LTBI during the study must repeat study assessments after TB prophylactic therapy has been received for at least 4 weeks. The Investigator and Medical Monitor and/or UCB study physician will decide which investigations (safety laboratory parameters, physical exams and questionnaires) need to be performed after required LTBI prophylaxis period and before the IMP will be restarted.

The IMP can be restarted no sooner than 4 weeks after the start of TB prophylactic therapy if it is deemed likely that the TB prophylactic therapy will be continued to full completion. If no TB prophylactic therapy is initiated for the newly diagnosed LTBI, the study participant must permanently stop IMP and be withdrawn from the study. Every related action should be discussed in advance with the Medical Monitor.

Study participants who prematurely discontinue treatment for LTBI or who, in the opinion of the Investigator or Sponsor, are noncompliant with anti-TB therapy must discontinue further intake of IMP and be immediately withdrawn. Once withdrawn from study treatment, study participants should return for the PEOT visit, complete all assessments, and complete the SFU visit. LTBI must be reported as an AE. Follow-up reports should be completed as per protocol requirement until such time as the LTBI resolves.

8.2.6.3.6 Active TB or non-tuberculosis mycobacterium infection

Study participants who develop active TB or NTM infection during the study must be withdrawn from the study. The study participant must be immediately permanently discontinued from study medication and a PEOT visit must be scheduled as soon as possible, but no later than the next scheduled visit. The study participant should be encouraged to keep the SFU visit as specified by the protocol. Treatment for active TB or NTMB should be started immediately.

Confirmed active TB is always considered an SAE. UCB's process requires that these must be captured on an SAE report form and provided to the Sponsor in accordance with SAE reporting requirements. Follow-up reports should be completed as per protocol requirement until such time as the TB infection resolves.

8.2.6.3.7 Tuberculosis management of LTBI, active TB, or other NTB infection identified during study

During the study, study participants who develop evidence of LTBI, active TB or NTB infection must immediately stop further administration of IMP and will be referred to a TB specialist (pulmonologist or infectious disease specialist) for further evaluation. Study participants diagnosed with active TB or LTBI should receive appropriate TB or prophylaxis therapy. If a TB specialist excludes active TB, the study participant can restart the IMP no earlier than 4 weeks after the start of an appropriate TB prophylactic therapy. The study participant should be transferred to the care of his or her physician and managed according to the standard of care.

Study participants identified as having active TB during the study must be withdrawn and scheduled to return for the PEOT Visit as soon as possible but no later than the next scheduled study visit and complete all PEOT Visit assessments. The study participant should be encouraged to complete an SFU Visit after the last dose of study medication.

If infection with NTM is identified during the study, the same procedure as for active TB acquired during the study and compliant TB treatment shall be followed.

Study participant eligibility, retesting requirements, and treatment requirements are shown in [Figure 8-1](#) (screening) and [Figure 8-2](#) (during the study). Additional details on TB detection and management are provided in the UCB TB Detection Procedure Guideline.

Figure 8-1: Decision tree for IGRA TB results at Screening

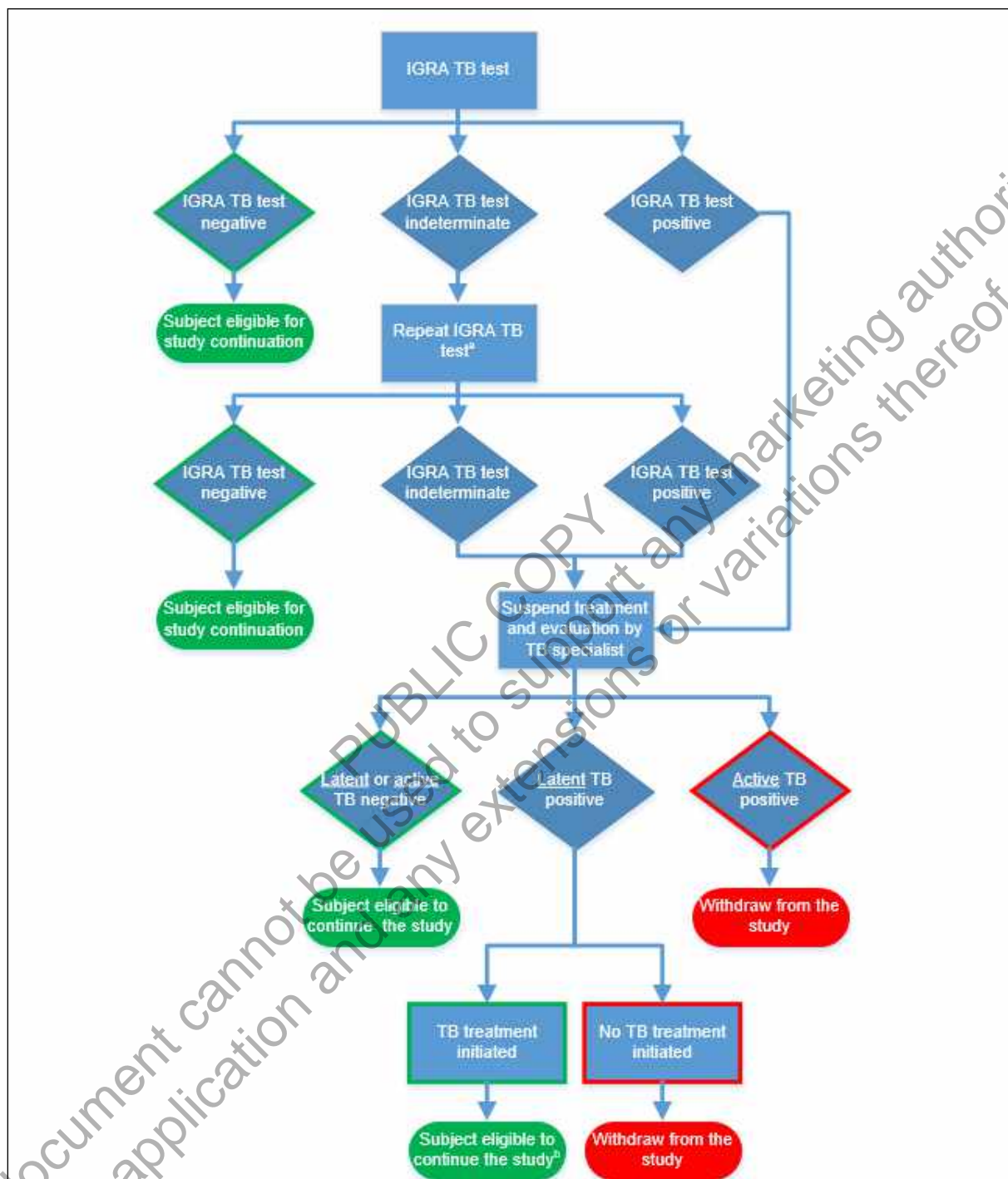


IGRA=interferon gamma release assay; IMP=investigational medicinal product; LTBI=latent tuberculosis infection; TB=tuberculosis

^a IGRA retest must be done during the protocol-defined Screening window

^b Study participants with LTBI diagnosed during Screening must have completed a course of prophylaxis prior to IMP dosing. Participants can be rescreened after completion of a course of prophylaxis plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline.

Figure 8-2: Decision tree for IGRA TB results during a study



ASAP=as soon as possible; IGRA=interferon-gamma release assay; IMP=investigational medicinal product; LTBI=latent tuberculosis infection; TB=tuberculosis

^a IGRA retest must be done ASAP and prior to the next IMP dose

^b Study participants with LTBI diagnosed during the study may continue the study only after they have completed at least 4 weeks of appropriate prophylactic therapy and thereafter, will continue and complete the entire regimen.

8.3 Adverse events and serious adverse events

AE will be reported by the study participant (or, when appropriate, by a caregiver, surrogate, or the study participant's legally authorized representative).

The Investigator and any designees are responsible for detecting, documenting, and recording events that meet the definition of an AE or SAE and remain responsible for following up AEs that are serious, considered related to the study treatment or study procedures, or that caused the study participant to discontinue the study treatment or the study (see Section 7).

Confirmed and suspected cases of COVID-19 infection will be recorded as AEs (or SAEs, as required).

8.3.1 Time period and frequency for collecting AE and SAE information

All AEs and SAEs will be collected from the signing of the ICF until the SFU visit (except for those study participants who enroll in extension study HS0005) or until the first dose administration in extension study HS0005 (for study participants enrolling in HS0005).

In order to ensure complete safety data collection, all AEs occurring during the study (ie, after the signing of the ICF), including any pretreatment and posttreatment periods required by the protocol, must be reported in the CRF even if no IMP was taken but specific study procedures were conducted. This includes all AEs not present prior to the initial visit and all AEs that recurred or worsened after the initial visit.

All SAEs will be recorded and reported to the Sponsor or designee within 24 hours, as indicated in Appendix 3. The Investigator will submit any updated SAE data to the Sponsor within 24 hours of it being available.

The Investigator is specifically requested to collect and report to UCB (or its representative) any SAEs (even if the Investigator is certain that they are in no way associated with the IMP), up to 20 weeks from the last dose of IMP for each study participant, and to also inform study participants of the need to inform the Investigator of any SAE within this period. Serious AEs that the Investigator thinks may be associated with the IMP must be reported to UCB regardless of the time between the event and the end of the study.

The method of recording, evaluating, and assessing causality of AE and SAE and the procedures for completing and transmitting SAE reports are provided in Section 10.3.

8.3.2 Method of detecting AEs and SAEs

Care will be taken not to introduce bias when detecting AEs and/or SAEs. Open-ended and non-leading verbal questioning of the participant is the preferred method to inquire about AE occurrences.

8.3.3 Follow-up of AEs and SAEs

After the initial AE/SAE report, the Investigator is required to proactively follow each study participant at subsequent visits/contacts. All AEs and SAEs, will be followed until resolution, stabilization, the Investigator determines that it is no longer clinically significant, the event is otherwise explained, or the study participant is lost to follow-up (as defined in Section 7.3). Further information on follow-up procedures is given in Section 10.3.

8.3.4 Regulatory reporting requirements for SAEs

Prompt notification by the Investigator to the Sponsor of a SAE is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a study treatment under clinical investigation are met.

The Sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study treatment under clinical investigation. The Sponsor will comply with country-specific regulatory requirements relating to safety reporting to the regulatory authority, IRB/IEC, and Investigators.

Investigator safety reports must be prepared for suspected unexpected serious adverse reactions (SUSAR) according to local regulatory requirements and Sponsor policy and forwarded to Investigators as necessary.

An Investigator who receives an Investigator safety report describing a SAE or other specific safety information (eg, summary or listing of SAEs) from the Sponsor will review and then file it along with the Investigator's Brochure and will notify the IRB/IEC, if appropriate according to local requirements.

8.3.5 Pregnancy

Details of all pregnancies in female participants and female partners of male participants who become pregnant will be collected after the start of study treatment and through the SFU visit (ie, 20 weeks after last dose of IMP).

If a pregnancy is reported, the Investigator must immediately inform the Sponsor within 24 hours of learning of the pregnancy and should follow the procedures outlined in Appendix 4.

A female study participant should be withdrawn from the study as soon as pregnancy is known (by positive pregnancy test), and the following should be completed:

- The study participant should return for an early discontinuation visit.
- The study participant should immediately stop the intake of the IMP.
- An SFU Visit should be scheduled 20 weeks after the study participant has received her last dose of IMP.

Abnormal pregnancy outcomes (eg, spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) are considered SAEs.

8.3.6 Adverse events of special interest

An AE of special interest is any AE that a regulatory authority has mandated be reported on an expedited basis, regardless of the seriousness, expectedness, or relatedness of the AE to the administration of a UCB product/compound. Potential Hy's Law cases, defined as $\geq 3 \times \text{ULN}$ ALT or AST with coexisting $\geq 2 \times \text{ULN}$ total bilirubin in the absence of $\geq 2 \times \text{ULN}$ ALP, with no alternative explanation for the biochemical abnormality, must ALWAYS be reported to UCB as an AE of special interest (ie, without waiting for any additional etiologic investigations to have been concluded). Follow-up information should then be reported if an alternative etiology is identified during investigation and monitoring of the participant.

8.3.7 Other safety topics of interest

Prespecified safety topics of interest for the study are infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behavior, major adverse cardiovascular events, hepatic events and PDILI, malignancies, and inflammatory bowel disease.

These are based on findings from the IMP clinical program to date, potential risks generally associated with biologic immunomodulators, or findings from other medicines with a related mechanism of action. There are no specific AE reporting requirements for these topics, except those listed below for events relating to TB; however, special monitoring, additional data collection activities, and/or enhanced signal detection activities (within UCB) are in place.

The reporting requirements for events relating to TB are as follows:

- The IGRA test conversions defined as a positive or indeterminate (and confirmed indeterminate on repeat) should be reported as AEs. The AE term would need to be updated with final diagnosis once available.
- Latent TB infection must be reported as an AE. Follow-up reports should be completed as per protocol requirement until the LTBI resolves.
- Confirmed active TB is always considered an SAE and must be reported per SAE reporting instruction in the study protocol. Follow-up reports should be completed as per protocol requirement until TB infection resolves.

8.3.8 Anticipated serious adverse events

The following list of Anticipated SAEs (Table 8-1) is predicted to occur in the population studied in this protocol at some frequency that is independent of drug exposure. This original list will remain in effect for the duration of the protocol. Note that listed events will not be regarded as anticipated SAEs if they are life threatening or if they result in the death of the study participant.

This list does not change the Investigator's obligation to report all SAEs (including Anticipated SAEs) as detailed in Section 8.3.1 and Section 10.3.

Table 8–1: Anticipated SAEs for the Population of Participants with HS

MedDRA System Organ Class	MedDRA Preferred Term
Gastrointestinal disorders	Crohn's disease Colitis ulcerative
Psychiatric disorders	Depression Anxiety
Musculoskeletal and connective tissue disorders	Arthropathy
Skin and subcutaneous tissue disorders	Pyoderma gangrenosum Pilonidal cyst Acne conglobate Hidradenitis
Neoplasms benign, malignant and unspecified (including cysts and polyps)	Lymphoma Squamous cell carcinoma of skin
Infections and infestations	Cellulitis
Metabolism and nutritional disorders	Diabetes mellitus Dyslipidaemia Metabolic syndrome
Endocrine disorders	Thyroid disorder Polycystic ovaries

HS=hidradenitis suppurativa; MedDRA=Medical Dictionary for Regulatory Activities; SAE=serious adverse event

8.3.9 Suspected transmission of an infectious agent

For the purposes of reporting, any suspected transmission of an infectious agent via a medicinal product should be considered as an SAE; such cases must be reported immediately, recorded in the AE module of the eCRF, and followed as any other SAE. Any organism, virus, or infectious particle (eg, prion protein transmitting transmissible spongiform encephalopathy), pathogenic or nonpathogenic, is considered an infectious agent.

8.4 Safety signal detection

Selected data from this study will be reviewed periodically to detect as early as possible any safety concern(s) related to the IMP so that Investigators, clinical study participants, regulatory authorities, and IRBs/IECs will be informed appropriately and as early as possible.

The UCB study physician or medically qualified designee/equivalent will conduct an ongoing review of SAEs and perform ongoing SAE reconciliations in collaboration with the Patient Safety (PS) representative.

As appropriate for the stage of development and accumulated experience with the IMP, medically qualified personnel at UCB may identify additional safety measures (eg, AEs, vital signs, laboratory or ECG results) for which data will be periodically reviewed during the course of the study.

An independent DMC will be used in this study; see Section 9.7 and Section 10.1.5 for details.

8.5 Treatment of overdose

For this study, any dose of IMP greater than that prescribed in the protocol will be considered an overdose. Overdose events are only considered AEs or SAEs if there are associated clinical signs and symptoms or if the act of taking the excess IMP itself is an AE or SAE (eg, suicide attempt).

UCB does not recommend specific treatment for an overdose. Any signs or symptoms of adverse reactions should be treated symptomatically as per standard care by the Investigator.

Bimekizumab will not be self-administered by the study participant.

In the event of an overdose, the Investigator should:

1. Contact the Medical Monitor immediately.
2. Closely monitor the study participant for any AE/SAE and laboratory abnormalities until they have resolved, have a stable sequelae, the Investigator determines that it is no longer clinically significant, or the study participant is lost to follow up.
3. Document the quantity of the excess dose as well as the duration of the overdose in the CRF.

Decisions regarding dose interruptions will be made by the Investigator in consultation with the Medical Monitor based on the clinical evaluation of the study participant.

8.6 Pharmacokinetics

Blood samples will be collected prior to dosing for measurement of plasma concentrations of bimekizumab at all timepoints described in Table 1–1. A total of 9mL will be collected at timepoints for which PK, anti-drug antibodies (ADAb), and neutralizing antibodies are all measured and 3mL will be collected at timepoints for which only PK is measured. Instructions for the collection and handling of biological samples will be provided by the Sponsor. The actual date and time (24-hour clock time) of each sample will be recorded.

Samples will be used to evaluate the PK of bimekizumab. Samples collected for analyses of bimekizumab plasma concentration may also be used to evaluate safety or efficacy aspects related to concerns arising during or after the study.

Genetic analyses will not be performed on these samples. Study participant confidentiality will be maintained. At visits during which blood samples for the determination of plasma concentrations of bimekizumab will be taken, 1 sample of sufficient volume can be used.

Drug concentration information that may unblind the study will not be reported to any blinded study personnel as long as the study remains blinded.

8.7 Genetics

For individuals consenting to the pharmacogenetic substudy, blood samples will be drawn for exploratory genetic/epigenetic analyses at the timepoints specified in Table 1–1. Collection of these samples will enable evaluation of genetics/epigenetics biomarkers relative to disease biology and progression, drug treatment and inflammatory and immune response processes. The nature and format of these tentative analyses will be determined when the results of the main study are made available. A separate ICF will be required for those study participants who agree to participate in the pharmacogenetics substudy. The substudy will be conducted only where

ethically accepted and authorized by the regulatory agencies. Refusal to participate in the substudy will not affect a study participant's ability to participate in the main study.

The samples will be stored at -80°C at the central biorepository for up to 20 years.

In the event of deoxyribonucleic acid (DNA) extraction failure, a replacement genetic blood sample may be requested from the participant. Signed informed consent will be required to obtain a replacement sample unless it was included in the original consent.

See Appendix 5 (Section 10.5) for information regarding genetic research. Details on processes for collection and shipment and destruction of these samples can be found in laboratory manual for this study. Detailed information on sample analysis will be provided in a bioanalytical report.

8.8 Pharmacodynamics

See Section 8.9.

8.9 Biomarkers

Where local regulations permit, blood samples will be drawn for exploratory ribonucleic acid, proteins and metabolites biomarker analysis at the timepoints specified in Table 1–1. Where local regulations permit, urine samples will be drawn for exploratory proteins and metabolites biomarker analysis at the timepoints specified in Table 1–1. Collection of these samples will enable evaluation of biomarkers relative to disease biology and progression, drug treatment and inflammatory and immune response processes.

Instructions pertaining to sample collection, processing, storage, labeling, and shipping are provided in the laboratory manual for this study. The nature and format of these tentative analyses will be determined at a later stage. The samples will be stored at the secure long-term storage facility selected by UCB for up to 20 years.

These samples will only be used to further our understanding of HS and/or how biomarkers, including genetic variation, may affect response or be affected by treatment with bimekizumab in HS.

8.9.1 Immunogenicity assessments

Blood samples for the measurement of ADA_b and neutralizing antibodies will be collected. A total of 9mL will be collected at timepoints for which PK, ADA_b, and neutralizing antibodies are all measured. Immunogenicity data will not be sent to the Investigator to protect the blinded nature of the treatment assignments and response.

Antibodies to bimekizumab will be evaluated in plasma samples collected from all participants according to the Schedule of Activities. Additionally, plasma samples should also be collected at the final visit from participants who discontinued IMP or were withdrawn from the study. These samples will be tested by the Sponsor or Sponsor's designee.

Plasma samples will be screened for antibodies binding to bimekizumab and the titer of confirmed positive samples will be reported. Other analyses may be performed to verify the stability of antibodies to bimekizumab and/or further characterize the immunogenicity of bimekizumab.

The detection and characterization of antibodies to bimekizumab will be performed using validated assay methods by or under the supervision of the Sponsor. All samples collected for detection of antibodies to bimekizumab will also be evaluated for bimekizumab plasma concentration to enable interpretation of the antibody data. Antibodies may be further characterized and/or evaluated for their ability to neutralize the activity of bimekizumab. Samples may be stored for a maximum of 20 years (or according to local regulations) following the last study participant's last visit for the study at a facility selected by the Sponsor to enable further analysis of immune responses to bimekizumab.

8.10 Medical resource utilization and health economics

Health-related outcomes and medical resource utilization will be collected as part of standard eCRF pages during the study (eg, concurrent medical procedures, concomitant medications, hospitalizations, WPAI-SHP).

8.11 Photography

At certain sites, where feasible, representative photographs of the changes in skin will be captured. Photographs will be anonymized. This is optional for study participants and requires a separate informed consent. A decision not to consent does not exclude the study participant from the study.

9 STATISTICAL CONSIDERATIONS

A description of statistical methods follows and will be described in more detail in the Statistical Analysis Plan (SAP).

9.1 Definition of analysis sets

9.1.1 Enrolled Set

The Enrolled Set will consist of all study participants who have given informed consent.

9.1.2 Randomized Set

The Randomized Set (RS) will consist of all randomized study participants.

9.1.3 Safety Set

The Safety Set will consist of all study participants who received at least 1 full or partial dose of IMP and will be used for the demographic, safety, and immunogenicity analyses.

9.1.4 Full Analysis Set

The Full Analysis Set (FAS) will consist of all study participants who received at least 1 dose (full or partial) of IMP and had a valid Baseline measurement and a post-Baseline measurement for abscess, inflammatory nodules, and draining tunnel counts.

9.1.5 Per-Protocol Set

The Per-Protocol Set will consist of all study participants in the FAS who had no important protocol deviations affecting the primary efficacy variable. Important protocol deviations will be predefined and study participants with important protocol deviations will be evaluated during ongoing data cleaning meetings prior to unblinding of the data.

9.1.6 Pharmacokinetics Per-Protocol Set

The Pharmacokinetics Per-Protocol Set will consist of study participants who received at least 1 full dose of bimekizumab and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the PK. The Pharmacokinetics Per-Protocol Set is defined separately for each of the treatment periods (ie, separately for the Initial Treatment Period and the Maintenance Treatment Period).

9.1.7 COVID-19 Free Set

The COVID-19 Free Set will consist of all study participants randomized into the study and who have no COVID-19 impact through Week 16. This analysis set will be used for sensitivity analysis of the primary efficacy endpoint.

9.2 General statistical considerations

All analyses will be performed using SAS® version 9.4 or later (SAS Institute, Cary, NC, US).

Descriptive statistics will be used to provide an overview of the Baseline, efficacy, and safety results. For categorical parameters, the number and percentage of study participants in each category will be presented by treatment group. The denominator for the percentages will be based on the number of study participants appropriate for the purpose of analysis. Unless otherwise noted, all percentages will be expressed to 1 decimal place. For continuous parameters, descriptive statistics will include n, mean, standard deviation, median, minimum, and maximum. Two-sided 95% confidence intervals, geometric means, and coefficient of variation will be presented for selected variables as appropriate.

Baseline for each assessment is defined as either the value obtained at Baseline or the last available value obtained prior to treatment administration (details to be specified in the SAP).

Formal statistical testing will be conducted for this study for the primary and secondary efficacy variables. Other efficacy variables will be summarized descriptively by treatment arm. P-values and confidence intervals may be produced for other or exploratory variables but will be interpreted as non-inferential (ie, nominal). Additionally, other analyses will be conducted as deemed appropriate and described in the SAP.

The primary treatment comparison for all formal statistical analyses of efficacy will be between bimekizumab and placebo.

9.3 Planned efficacy/outcome analyses

9.3.1 Analysis of the primary efficacy endpoint

The primary objective of this randomized, double-blind, placebo-controlled, multicenter, pivotal study in study participants with moderate to severe HS is to compare the efficacy of bimekizumab 320mg Q2W and bimekizumab 320mg Q4W with placebo at Week 16. For the purposes of Week 16 analyses, the bimekizumab treatment arms of 320mg Q2W/Q2W and bimekizumab 320mg Q2W/Q4W treatment groups will be pooled.

The primary and secondary efficacy analyses will be performed based on the RS.

The primary endpoint is the HiSCR₅₀ response at Week 16. The primary efficacy analysis will evaluate the composite estimand in the RS. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and early receipt of

systemic antibiotic rescue medication, or discontinuation of IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for the primary efficacy analysis:

4. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
5. Study participant-level outcome=HiSCR₅₀ at Week 16.
6. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, and not discontinuing study treatment due to an AE or lack of efficacy through Week 16. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation.
7. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (ie, receipt of systemic antibiotic rescue medication, or discontinuation of IMP due to an AE or lack of efficacy) will be imputed as non-response.

The statistical hypothesis for the HiSCR₅₀ response at Week 16 is that the conditional odds ratio for the HiSCR₅₀ response in the bimekizumab treatment group relative to the placebo group is equal to 1.

A logistic regression model will be used to assess the effect of bimekizumab vs placebo on HiSCR₅₀ response. The model will include a fixed effect for treatment. The stratification variables of Hurley stage and prior antibiotic use will be added to the model unless inappropriate. The odds ratio versus placebo, p-value (from Wald test), and confidence interval will be calculated.

To assess the impact of the COVID-19 pandemic on the primary efficacy endpoint analysis, additional sensitivity analyses will be performed as specified in Section 9.3.4.

9.3.2 Analysis of the secondary efficacy endpoints

The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment:

1. Proportion of study participants who achieve HiSCR₇₅ at Week 16. A similar analysis to the primary responder analysis will be performed.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
2. Absolute change from Baseline in DLQI Total Score at Week 16. Analysis will be based on an analysis of covariance (ANCOVA) with treatment and stratification variables as fixed effects and the Baseline values as covariate.
 - a. bimekizumab 320mg Q2W vs placebo

- b. bimekizumab 320mg Q4W vs placebo
- 3. Absolute change from Baseline in Skin Pain Score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HSSDD. Analysis will be based on an ANCOVA with treatment and stratification variables as fixed effects and the Baseline values as covariate.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
- 4. Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) at Week 16. A similar analysis to the primary responder analysis will be performed.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo

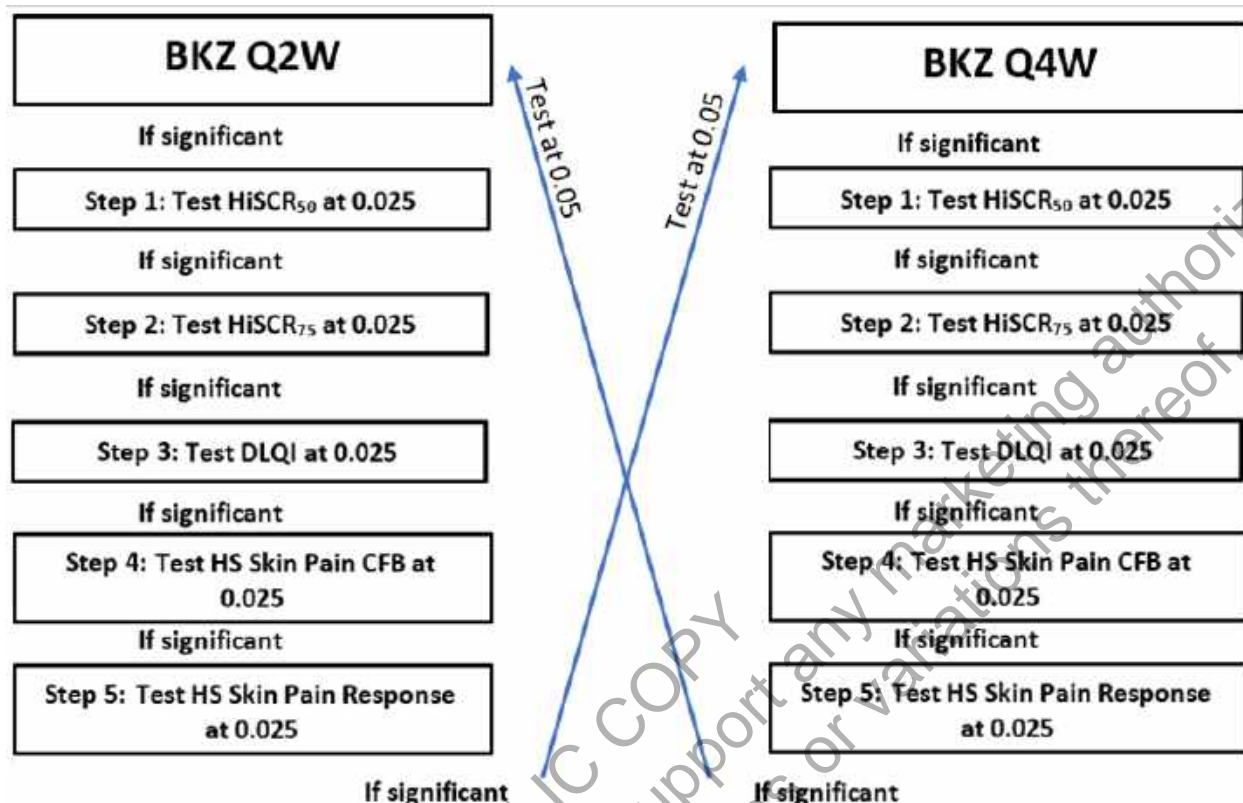
To control the overall type I error rate at 0.05 for the multiple comparisons in the primary and secondary efficacy variables, a closed testing procedure under a parallel gatekeeping framework will be applied (Sun, 2018).

Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$). Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

1. Step 1: Test HiSCR₅₀ at significance level 0.025.
2. Steps 2 to 5 – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown below, moving to the next step only if significance achieved at 0.025.
3. In the event that Step 5 is significant at 0.025 for a given dose, then Steps 1 to 5 will be repeated for the other dose using a significance level of 0.05.

A schematic of the procedure is shown in Figure 9–1.

Figure 9–1: Closed Testing Procedure



AN=abscess and inflammatory nodule; CFB=change from Baseline; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks

9.3.3 Other efficacy/other outcome analyses

Analyses of the other efficacy measures will be detailed in the SAP.

9.3.4 COVID-19 impact analysis

To assess the impact of the COVID-19 pandemic on the primary efficacy endpoint analysis, the following additional sensitivity analyses will be performed:

- Imputation as nonresponse for any missing data due to COVID-19 at Week 16 and analyzed using the same analysis model as for the primary analysis
- Separate inferential analysis of the primary and secondary efficacy endpoints based on the COVID-19 Free Set
- Separate summary statistics for the primary efficacy endpoint at Week 16, based on the COVID-19 Free Set
- Summary of the number of study participants with missing primary or secondary endpoint data, as applicable

In addition, to assess the broader impact of the COVID-19 pandemic on the study, the following summaries will be presented:

- Summary of study participant disposition based on enrollment before, during, and after COVID-19 pandemic onset
- A summary of study visits impacted by the COVID-19 pandemic
- A summary of protocol deviations related to COVID-19

Any additional COVID-19 related analyses will be specified in the study SAP. Note: the date of COVID-19 pandemic onset is defined as 11 March 2020, the date that the World Health Organization declared the COVID-19 pandemic. The end date of the COVID-19 pandemic end will be defined in the study SAP, if applicable.

9.4 Planned safety and other analyses

9.4.1 Safety analyses

All TEAEs, SAEs, TEAEs leading to discontinuation, AEs of special interest (eg, cases meeting Hy's Law criteria), and other safety topics of interest (Section 8.3.7) will be collected during the study and for up to 20 weeks after the last dose of IMP (for study participants who do not participate in the extension study, HS0005). Safety analyses will be carried out using the Safety Set (study participants who received at least 1 full or partial dose of IMP). Summaries of Confirmed and Suspected COVID-19 TEAEs, respectively, will be presented. The definition of Confirmed and Suspected COVID-19 TEAEs will be provided in the SAP.

9.4.2 PK and ADAb analyses

Plasma concentrations of bimekizumab will be summarized by treatment group at each timepoint using descriptive statistics. In addition, PK model-based analyses may be performed.

Antidrug antibody data will be evaluated for each study participant and each regimen, and rates and classification of ADAb-positive study participants will be calculated.

9.5 Handling of protocol deviations

Important protocol deviations are identified as part of the data cleaning process as specified in the study-specific data cleaning schedule/plan. Ongoing, blinded data cleaning meetings will be held throughout the duration of the study. Objectives of these meetings include to review protocol deviations and to document potential impact that these deviations might have on the study objectives. Furthermore, overall trends in protocol deviations will be discussed at the Data Evaluation Meeting(s). Through this ongoing data cleaning and evaluation process, all decisions regarding important protocol deviations are made on an ongoing basis. Protocol deviations (eg, missing assessments or visits) related to COVID-19 will be listed and summarized separately.

9.6 Handling of dropouts or missing data

The analyses for the primary and secondary efficacy variables will include the use of multiple imputation. In multiple imputation, the missing value is replaced by a set of plausible values, where each value is a Bayesian draw from the conditional distribution of the missing data given the observed data.

Intermittent missing data will be imputed using the Markov-Chain Monte Carlo method, followed by regression for monotone missing data. The multiple imputation procedures planned for the primary and secondary efficacy analyses are based on an assumption of data missing at random.

The sensitivity of results to the approach for handling missing data will be evaluated via supportive analyses using different missing data mechanisms. The following supportive analyses for the primary efficacy variable will be conducted:

1. Deviations from the missing at random pattern will be evaluated using a reference-based multiple imputation approach. Intermittent missing data will be imputed using the Markov-Chain Monte Carlo method. The remaining monotone missing data will be assumed to follow a missing not at random pattern. These data will be imputed using a reference-based approach in which the multiple imputation model is based on data from the placebo group, thereby assuming that monotone missing data follow a trajectory similar to the placebo group.
2. Tipping point analyses will be performed to evaluate missingness assumptions. Various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. It will include scenarios where study participants who have missing data and are randomized to bimekizumab have a lower probability of response compared to study participants who have missing data and were randomized to placebo. For binary variables, this includes the worst case scenario where study participants who have missing data and are randomized to bimekizumab are considered nonresponders, while study participants who have missing data and were randomized to placebo are considered responders. The goal of the tipping point analysis is to systematically vary assumptions until there is no evidence of a treatment effect (if possible). The plausibility of such required delta adjustment will then be assessed.
3. The treatment policy strategy for addressing intercurrent events will be considered. This will be based on an analysis of all available data at Week 16 regardless of the occurrence of intercurrent events. This analysis will use the same models specified for the primary analysis, where study participants are analyzed according to their randomized treatment, even if they discontinued study treatment for any reason prior to Week 16. Even though efforts will be made to collect the primary outcome data for all study participants at Week 16, there may still be some study participants for whom Week 16 efficacy data cannot be obtained. In this case, missing data will be imputed using multiple imputation under the assumption of MAR. Results will be combined into a single inference using Rubin's rule. It should be noted that this measures something different from the primary analysis and could be confounded by placebo study participants who withdraw and are subsequently on another active medication at the time of the Week 16 assessment. Therefore, the results of this analysis should be interpreted in the appropriate context.
4. An additional supportive analysis will be based on observed data only for study participants who are still on the randomized treatment at Week 16. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing. The same procedure described as in the primary efficacy analyses will be used.

All imputation of other missing data will be detailed in the SAP.

9.7 Planned interim analysis and data monitoring

9.7.1 Data Monitoring Committee

An independent DMC will periodically review unblinded efficacy and safety data to assess the benefit/risk of bimekizumab in study participants with moderate to severe HS. Individual study participant-level efficacy data listings may be provided to the DMC to put the safety review in the context of risk/benefit. Any data to be provided will be specified per the DMC charter. ■

9.7.2 Interim Analysis

After the last study participant has completed the Week 48 visit, or after the last study participant has prematurely discontinued prior to reaching Week 48, an unblinded interim analysis will be performed and a corresponding interim clinical study report (CSR) may be written. The purpose of this interim analysis is to perform a comprehensive evaluation of all available double-blind data to prepare regulatory submissions for approval of the HS indication. A final analysis and updated final CSR will be prepared once all data (through to the SFU visit) have been collected.

9.8 Determination of sample size

A total of 490 study participants will be randomly assigned in a 2:2:2:1 ratio to the following treatment arms:

- Bimekizumab 320mg Q2W during Initial Treatment Period (Weeks 0-16) and Maintenance Treatment (Weeks 16-48) Periods, N=140
- Bimekizumab 320mg Q2W during Initial Treatment Period (Weeks 0-16), and Bimekizumab 320mg Q4W during Maintenance Treatment Period (Weeks 16-48), N=140
- Bimekizumab 320mg Q4W during Initial Treatment (Weeks 0-16) and Maintenance Treatment Periods (Weeks 16-48), N=140
- Placebo during Initial Treatment Period (Weeks 0-16), and Bimekizumab 320mg Q2W during Maintenance Treatment Period (Weeks 16-48), N=70

The analysis of the primary efficacy endpoint and secondary efficacy endpoints are based on a comparison of bimekizumab versus placebo at Week 16, with alpha adjustment strategy as indicated in Section 9.3.1 and Section 9.3.2.

The power to detect a statistically significant difference for each of the endpoints are shown in [Table 9–1](#). Notably, with a 2-sided significance level of 0.025, the sample size of 140:70

provides 73% power for detecting at least a difference of 1.5 (bimekizumab Q4W vs placebo) for the Worst Pain change from Baseline (CFB) endpoint.

Given the high level of power for each of the primary and key secondary endpoints at the 0.025 significance level for the Q2W comparison (power ≥ 0.89), and per the alpha spending strategy, there is a high likelihood that the Q4W comparison of Worst Pain CFB vs placebo will be allowed to be tested against the 0.05 level of significance. The power for this latter test is 81%. The sample size is thus ultimately driven by the Worst CFB endpoint. Furthermore, the randomization ratio of 2:2:2:1 has been chosen to provide study participants with a high probability (6/7 ~ 86%) of being randomized to active study drug.

After randomization for this study was complete, an additional endpoint to assess Worst Pain response was included in the sequential testing procedure. This additional endpoint is defined as HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change at Week 16. Note that the power calculations reported in [Table 9-1](#) for this endpoint are based on the sample size that was initially driven by the Worst Pain CFB endpoint as described above. With a 2-sided significance level of 0.025, the sample size of 104:52 in the subset of participants reporting Baseline HSSDD worst skin pain score at or beyond the threshold for clinically meaningful change (ie, Baseline HSSDD \geq threshold value) provides 53% power for detecting a statistically significant difference between bimekizumab Q4W and placebo in the proportion of Worst Pain responders.

Note that the power at the 0.025 level of significance associated with this endpoint for the comparison between bimekizumab Q2W and placebo is 95%. The Q4W comparison of Worst Pain response vs placebo against the 0.05 level of significance is therefore likely, and the power at this significance level is 65%. Given the strength of the power for the bimekizumab Q2W arm vs placebo across endpoints, it is considered acceptable to have a relatively low power for this final endpoint in the testing sequence for the bimekizumab Q4W treatment arm vs placebo.

Table 9–1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6
Worst Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0003 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms

^b Within-subject average of Worst Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Pain score at or above the threshold for clinically meaningful change from Baseline

10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

10.1 Appendix 1: Regulatory, ethical, and study oversight considerations

10.1.1 Regulatory and ethical considerations

The study will be conducted under the auspices of an IRB/IEC, as defined in local regulations, International Council for Harmonisation (ICH)- Good Clinical Practice (GCP), and in accordance with the ethical principles that have their origin in the Declaration of Helsinki.

The Investigator/UCB will ensure that an appropriately constituted IRB/IEC that complies with the requirements of the current ICH-GCP version or applicable country-specific regulations will be responsible for the initial and continuing review and approval of the clinical study. Prior to initiation of the study, the Investigator/UCB will forward copies of the protocol, ICF, Investigator's Brochure, Investigator's curriculum vitae (if applicable), advertisement (if applicable), and all other participant-related documents to be used for the study to the IRB/IEC for its review and approval.

Before initiating a study, the Investigator will have written and dated full approval from the responsible IRB/IEC for the protocol.

The Investigator will also promptly report to the IRB/IEC all changes in the study, all unanticipated problems involving risks to participants or others, and any protocol deviations, to eliminate immediate hazards to participants.

The Investigator will not make any changes in the study or study conduct without IRB/IEC approval, except where necessary to eliminate apparent immediate hazards to the participants. For minor changes to a previously approved protocol during the period covered by the original approval, it may be possible for the Investigator to obtain an expedited review by the IRB/IEC as allowed.

As part of the IRB/IEC requirements for continuing review of approved studies, the Investigator will be responsible for submitting periodic progress reports to the IRB/IEC (based on IRB/IEC requirements), at intervals appropriate to the degree of participant risk involved, but no less than once per year. The Investigator should provide a final report to the IRB/IEC following study completion.

UCB (or its representative) will communicate safety information to the appropriate regulatory authorities and all active Investigators in accordance with applicable regulatory requirements. The appropriate IRB/IEC will also be informed by the Investigator or the Sponsor, as specified by the applicable regulatory requirements in each concerned country. Where applicable, Investigators are to provide the Sponsor (or its representative) with evidence of such IRB/IEC notification.

10.1.2 Financial disclosure

Insurance coverage will be handled according to local requirements.

Finance and insurance are addressed in the Investigator and/or CRO agreements, as applicable.

10.1.3 Informed consent process

Study participant's informed consent must be obtained and documented in accordance with local regulations, ICH-GCP requirements, and the ethical principles that have their origin in the principles of the Declaration of Helsinki.

Prior to obtaining informed consent, information should be given in a language and at a level of complexity understandable to the study participant in both oral and written form by the Investigator (or designee). Each study participant will have the opportunity to discuss the study and its alternatives with the Investigator.

Prior to participation in the study, the ICF should be signed and personally dated by the participant, or his or her legal representative, and by the person who conducted the informed consent discussion (Investigator or designee). The study participant or his or her legal representative must receive a copy of the signed and dated ICF. As part of the consent process, each study participant must consent to direct access to his or her medical records for study-related monitoring, auditing, IRB/IEC review, and regulatory inspection.

If the ICF is amended during the study, the Investigator (or the Sponsor, if applicable) must follow all applicable regulatory requirements pertaining to the approval of the amended ICF by the IRB/IEC and use of the amended form.

All studies conducted at centers in the US must include the use of a Health Insurance Portability and Accountability Act Authorization form.

The study participant may withdraw his or her consent to participate in the study at any time. A study participant is considered as enrolled in the study when he or she has signed the ICF. A CRF must not be started, nor may any study specific procedure be performed for a given study participant, without having obtained his or her written consent to participate in the study.

10.1.4 Data protection

UCB staff (or designee) will affirm and uphold the study participant's confidentiality. Throughout this study, all data forwarded to UCB (or designee) will be identified only by the study participant number assigned at Screening.

The Investigator agrees that representatives of UCB, its designee, representatives of the relevant IRB/IEC, or representatives of regulatory authorities will be allowed to review that portion of the study participant's primary medical records that directly concerns this study (including, but not limited to, laboratory test result reports, ECG reports, admission/discharge summaries for hospital admissions occurring during a study participant's study participation, and autopsy reports for deaths occurring during the study).

The study participant must be informed that his or her personal study-related data will be used by the Sponsor in accordance with local data protection law. The level of disclosure must also be explained to the study participant.

The study participant must be informed that his or her medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the Sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

10.1.5 Committees structure

A DMC will be reviewing safety and efficacy data on an ongoing basis. The DMC membership includes experienced clinicians and a statistician, all of whom have expertise in clinical studies. Further details will be specified in the DMC Charter.

Cardiovascular, Gastrointestinal, and Neuropsychiatric Adjudication Committees will also periodically review data from this study. Details will be provided in the Cardiovascular, Gastrointestinal, and Neuropsychiatric Adjudication Committee charters.

Both DMC and Adjudication Committee members may not participate in the study as principal or co-Investigators, or as study participant care physicians and must not be members of the study team at UCB or the conducting CRO. The duration of membership for the committees will be inclusive of planned analyses for this study.

10.1.6 Data quality assurance

All study participant data relating to the study will be recorded on printed or electronic CRF unless transmitted to the Sponsor or designee electronically (eg, laboratory data). The Investigator is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

The Investigator must maintain accurate documentation (source data) that supports the information entered in the CRF.

The Investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

The Sponsor or designee is responsible for the data management of this study including quality checking of the data.

Study monitors will perform ongoing source data verification to confirm that data entered into the CRF by authorized site personnel are accurate, legible, contemporaneous, original, and attributable from source documents; that the safety and rights of study participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH-GCP, and all applicable regulatory requirements. Some study-specific assessments may be conducted remotely according to the study protocol during a pandemic or other exceptional circumstance (eg, hurricanes) due to the need to implement safety measures and guidance from regulatory authorities (refer to Section 8). Remote monitoring visits may be conducted during the COVID-19 pandemic or under other exceptional circumstances as deemed appropriate to ensure study participants' safety where local regulations permit.

All essential documents must be retained by the Investigator for the minimum retention period mandatory under the applicable local laws and regulations. The Investigator will contact UCB for authorization prior to the destruction of any study records or in the event of accidental loss or destruction of any study records. The Investigator will also notify UCB should he or she relocate or move the study-related files to a location other than that specified in the Sponsor's trial master file.

Quality tolerance limits will be established for the study using parameters related to patient safety reporting and reliability of study results. The parameters will be monitored throughout the

study to identify systematic issues. Parameters used, parameter values, important deviations from the quality tolerance limits, and actions taken will be summarized in the clinical study report.

10.1.6.1 Case report form completion

The Investigator is responsible for prompt reporting of accurate, complete, and legible data in the electronic CRFs and in all required reports.

Any change or correction to the CRF after saving must be accompanied by a reason for the change.

Corrections made after the Investigator's review and approval (by means of a password/electronic signature) will be reapproved by the Investigator.

The Investigator should maintain a list of personnel authorized to enter data into the electronic CRF.

Detailed instructions will be provided in the CRF Completion Guidelines.

10.1.7 Source documents

All source documents must be accurate, clear, unambiguous, permanent, and capable of being audited. They should be made using some permanent form of recording (ink, typing, printing, optical disc). They should not be obscured by correction fluid or have temporary attachments (such as removable self-stick notes). Photocopies and/or printouts of CRFs are not considered acceptable source documents.

Source documents are original records in which raw data are first recorded. These may include hospital/clinic/general practitioner records, charts, diaries, x-rays, laboratory results, printouts, pharmacy records, care records, ECG or other printouts, completed scales, quality of life questionnaires, or video, for example. Source documents should be kept in a secure, limited access area.

Electronic Patient-Reported Outcome (ePRO) measures (eg, DLQI, EQ-5D-3L, WPAI-SHP, TSQM-9, HiSQOL, Daily HS Symptom Diary, HS Symptom Questionnaire, PGI-S-HS, PGI-C-HS, PGI-S-SP, and PGI-C-SP) will be completed by each participant and will be collected electronically.

The data collection and database management system will be supplied by a vendor and will be compliant with the relevant regulations. The data collected on the ePROs will be uploaded to a central server database and will be sent electronically to UCB (or a designated CRO).

Source documents that are computer generated and stored electronically must be printed for review by the monitor (eg, ECG reports). Once printed, these copies should be signed and dated by the Investigator and become a permanent part of the study participant's source documents. The Investigator will facilitate the process for enabling the monitor to compare the content of the printout and the data stored in the computer to ensure all data are consistent.

Electronic data records must be saved and stored as instructed by UCB (or designee).

10.1.8 Study and Site Closure

The Sponsor designee reserves the right to close the study site or terminate the study at any time for any reason at the sole discretion of the Sponsor. Study sites will be closed upon study

completion. A study site is considered closed when all required documents and study supplies have been collected and a study-site closure visit has been performed.

The Investigator may initiate study-site closure at any time, provided there is reasonable cause and sufficient notice is given in advance of the intended termination.

Reasons for the early closure of a study site by the Sponsor or Investigator may include but are not limited to:

- Failure of the Investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the Sponsor's procedures, or GCP guidelines
- Inadequate recruitment of participants by the Investigator
- Discontinuation of further bimekizumab development

10.1.9 Publication Policy

The results of this study may be published or presented at scientific meetings. If this is foreseen, the Investigator agrees to submit all manuscripts or abstracts to the Sponsor before submission. This allows the Sponsor to protect proprietary information and to provide comments.

The Sponsor will comply with the requirements for publication of study results. In accordance with standard editorial and ethical practice, the Sponsor will generally support publication of multicenter studies only in their entirety and not as individual site data. In this case, a coordinating Investigator will be designated by mutual agreement.

Authorship will be determined by mutual agreement and in line with International Committee of Medical Journal Editors authorship requirements.

10.2 Appendix 2: Clinical laboratory tests

- The tests detailed in the table below will be performed by the central laboratory.
- Protocol-specific requirements for inclusion or exclusion of participants are detailed in Section 5.1 and Section 5.2 of the protocol.
- Additional tests may be performed at any time during the study as determined necessary by the Investigator or required by local regulations.

Protocol-Required Safety Laboratory Assessments

Laboratory Assessments	Parameters			
Hematology	Platelet Count	<u>RBC Indices:</u> Mean corpuscular hemoglobin Mean corpuscular hemoglobin concentration Mean corpuscular volume		<u>White Blood Cell Count with Differential:</u> Neutrophils Lymphocytes Atypical lymphocytes Monocytes Eosinophils Basophils
	RBC Count			
	Hemoglobin			
	Hematocrit			
Clinical Chemistry	Blood urea nitrogen	Potassium	Alanine aminotransferase ^a	Total and direct bilirubin ^{a,b}
	Creatinine	Sodium	Aspartate aminotransferase ^a	Glucose (record fasting or nonfasting in CRF)
	Bicarbonate	Calcium	Alkaline phosphatase	Gamma glutamyltransferase
	Uric acid	Chloride	Magnesium	hs-CRP ^c
	Lactate dehydrogenase	Total cholesterol		
Routine Urinalysis	Specific gravity pH, glucose, protein, ketones, nitrite, blood by dipstick Microscopic examination (if blood or protein is abnormal)			
Other Screening Tests	Pregnancy testing ^d Follicle stimulating hormone ^e Urine drug screen (amphetamines, barbiturates, cocaine, opiates, cannabinoids, and benzodiazepines) Serology (human immunodeficiency virus, Hepatitis B, Hepatitis C)			

Protocol-Required Safety Laboratory Assessments

ALT=alanine aminotransferase; AST=aspartate aminotransferase; CRF=case report form;

hs-CRP=high-sensitivity C-reactive protein; RBC=red blood cell; ULN=upper limit of normal

^a Details of liver chemistry stopping criteria and required actions and follow-up assessments after liver stopping or monitoring event are given in Section 7.1.2 and Appendix 6 (Section 10.6). All events of $\geq 3 \times$ ULN ALT or AST with coexisting $\geq 2 \times$ ULN total bilirubin in the absence of $\geq 2 \times$ ULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality may indicate severe liver injury (possible Hy's Law) and must be reported as an SAE (excluding studies of hepatic impairment or cirrhosis).

^b If total bilirubin is $>$ ULN, a direct bilirubin estimation (%) will be performed.

^c hs-CRP will be tested at specified visits (Table 1-1).

^d A serum pregnancy test will be performed at Screening for all women of childbearing potential. A urine pregnancy test (urine dipstick analyzed locally) is also required at the Baseline Visit and all other visits in the Schedule of Activities (Table 1-1). Pregnancy test results must be negative prior to administering IMP.

^e A high follicle stimulating hormone level in the postmenopausal range may be used to confirm a postmenopausal state in women not using hormonal contraception or hormonal replacement therapy. However, in the absence of 12 months of amenorrhea, a single follicle stimulating hormone measurement is insufficient.

Investigators must document his or her review of each laboratory safety report.

Laboratory and/or analyte results (eg, hs-CRP, immunogenicity, PK) that could unblind the study will not be reported to any blinded study personnel as long as the study remains blinded.

10.3 Appendix 3: Adverse events – definitions and procedures for recording, evaluating, follow-up, and reporting

Definition of AE

AE Definition
<ul style="list-style-type: none">An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of IMP, whether or not considered related to the IMP.Note: An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of IMP.

Events Meeting the AE Definition
<ul style="list-style-type: none">Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (eg, ECG, radiological scans, vital signs measurements), including those that worsen from Baseline, considered clinically significant in the medical and scientific judgment of the Investigator (ie, not related to progression of underlying disease).Exacerbation of a chronic or intermittent pre-existing condition including either an increase in frequency and/or intensity of the condition.New conditions detected or diagnosed after IMP administration even though it may have been present before the start of the study.Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction.Signs, symptoms, or the clinical sequelae of a suspected overdose of either IMP or a concomitant medication. Overdose per se will not be reported as an AE/SAE unless it is an intentional overdose taken with possible suicidal/self-harming intent. Such overdoses should be reported regardless of sequelae."Lack of efficacy" or "failure of expected pharmacological action" per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments. However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as AE or SAE if they fulfil the definition of an AE or SAE.

Events NOT Meeting the AE Definition

- Any clinically significant abnormal laboratory findings or other abnormal safety assessments which are associated with the underlying disease, unless judged by the Investigator to be more severe than expected for the participant's condition.
- The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the study participant's condition.
- Medical or surgical procedure (eg, endoscopy, appendectomy): the condition that leads to the procedure is the AE.
- Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital).
- Anticipated day-to-day fluctuations of pre-existing disease(s) or condition(s) present or detected at the start of the study that do not worsen.

Definition of SAE

If an event is not an AE per definition above, then it cannot be an SAE even if serious conditions are met (eg, hospitalization for signs/symptoms of the disease under study, death due to progression of disease).

A SAE is defined as any untoward medical occurrence that, at any dose:
a. Results in death
b. Is life-threatening The term 'life-threatening' in the definition of 'serious' refers to an event in which the participant was at risk of death at the time of the event. It does not refer to an event, which hypothetically might have caused death, if it were more severe.
c. Requires inpatient hospitalization or prolongation of existing hospitalization In general, hospitalization signifies that the participant has been detained (usually involving at least an overnight stay) at the hospital or emergency ward for observation and/or treatment that would not have been appropriate in the physician's office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether "hospitalization" occurred or was necessary, the AE should be considered serious. Hospitalization for elective treatment of a pre-existing condition that did not worsen from Baseline is not considered an AE.
d. Results in persistent disability/incapacity <ul style="list-style-type: none">• The term disability means a substantial disruption of a person's ability to conduct normal life functions.• This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza, and accidental trauma (eg, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.
e. Is a congenital anomaly/birth defect
f. Important medical events: <ul style="list-style-type: none">• Medical or scientific judgment should be exercised in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.• Examples of such events include, but are not limited to, potential Hy's law, invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

Recording and Follow-Up of AE and/or SAE

AE and SAE Recording

- When an AE/SAE occurs, it is the responsibility of the Investigator to review all documentation (eg, hospital progress notes, laboratory reports, and diagnostics reports) related to the event.
- The Investigator will then record all relevant AE/SAE information in the CRF.
- It is **not** acceptable for the Investigator to send photocopies of the participant's medical records to UCB in lieu of completion of the AE/SAE CRF page.
- There may be instances when copies of medical records for certain cases are requested by UCB. In this case, all participant identifiers, with the exception of the participant number, will be redacted on the copies of the medical records before submission to UCB.
- The Investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. Whenever possible, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.

Assessment of Intensity

The Investigator will make an assessment of intensity for each AE and SAE reported during the study and assign it to 1 of the following categories:

- Mild: An event that is easily tolerated by the participant, causing minimal discomfort and not interfering with everyday activities.
- Moderate: An event that causes sufficient discomfort and interferes with normal everyday activities.
- Severe: An event that prevents normal everyday activities. An AE that is assessed as severe should not be confused with a SAE. Severe is a category utilized for rating the intensity of an event; and both AEs and SAEs can be assessed as severe.
- An event is defined as 'serious' when it meets at least 1 of the predefined outcomes as described in the definition of an SAE, NOT when it is rated as severe (eg, a severe AE may be either serious or not serious, depending on whether these criteria are also met).

The National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) should be used as a supportive standardization instrument to evaluate AEs and SAEs but the final intensity grading by the Investigator must be mild, moderate, or severe.

Assessment of Causality

- The Investigator is obligated to assess the relationship between IMP and each occurrence of each AE/SAE.
- A "reasonable possibility" of a relationship conveys that there are facts, evidence, and/or arguments to suggest a causal relationship, rather than a relationship cannot be ruled out.
- The Investigator will use clinical judgment to determine the relationship.
- Alternative causes, such as underlying disease(s), concomitant therapy, and other risk factors, as well as the temporal relationship of the event to IMP administration will be considered and investigated.
- The Investigator will also consult the Investigator's Brochure and/or Product Information, for marketed products, in his or her assessment.
- For each AE/SAE, the Investigator **must** document in the medical notes that he or she has reviewed the AE/SAE and has provided an assessment of causality.
- There may be situations in which an SAE has occurred and the Investigator has minimal information to include in the initial report to UCB. However, **it is very important that the Investigator always make an assessment of causality for every event before the initial transmission of the SAE data to UCB.**
- The Investigator may change his or her opinion of causality in light of follow-up information and send a SAE follow-up report with the updated causality assessment.
- The causality assessment is one of the criteria used when determining regulatory reporting requirements.

Follow-up of AEs and SAEs

- The Investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by UCB to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.
- An AE should be followed until it has resolved, has a stable sequelae, the Investigator determines that it is no longer clinically significant, or the participant is lost to follow up. This follow-up requirement applies to AEs, SAEs, and AEs of special interest.
- If a participant dies during participation in the study or during a recognized follow-up period, the Investigator will provide UCB with a copy of any post-mortem findings including histopathology.
- New or updated information will be recorded in the originally completed CRF.
- The Investigator will submit any updated SAE data to UCB within 24 hours of receipt of the information.

Reporting of SAEs

SAE Reporting to UCB via an Electronic Data Collection Tool

- The primary mechanism for reporting an SAE to UCB will be the electronic data collection tool.
- If the electronic system is unavailable for more than 24 hours, then the site will use the paper SAE data collection tool (see next section).
- The site will enter the SAE data into the electronic system as soon as it becomes available.
- After the study is completed at a given site, the electronic data collection tool will be taken off-line to prevent the entry of new data or changes to existing data.
- If a site receives a report of a new SAE from a study participant or receives updated data on a previously reported SAE after the electronic data collection tool has been taken off-line, then the site can report this information on a paper SAE form (see next section) or to the UCB study physician by telephone.
- Contacts for SAE reporting can be found on the page after the title page of this protocol.

SAE Reporting to UCB via Paper CRF

- Facsimile transmission of the SAE paper CRF is the preferred method to transmit this information to the Medical Monitor or UCB study physician.
- In rare circumstances and in the absence of facsimile equipment, notification by telephone is acceptable with a copy of the SAE data collection tool sent by overnight mail or courier service.
- Initial notification via telephone does not replace the need for the Investigator to complete and sign the SAE CRF pages within the designated reporting time frames.
- Contacts for SAE reporting can be found on the page after the title page of this protocol.

10.4 Appendix 4: Contraceptive guidance and collection of pregnancy information

Definitions

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile (see below).

Women in the following categories **are not considered WOCBP**:

1. Premenarchal
2. Premenopausal female with 1 of the following:
 - Documented hysterectomy
 - Documented bilateral salpingectomy
 - Documented bilateral oophorectomy

Note: Documentation can come from the site personnel's review of the participant's medical records, medical examination, or medical history interview.

3. Postmenopausal female
 - A postmenopausal state is defined as no menses for 12 months without an alternative medical cause. A high follicle stimulating hormone level in the postmenopausal range may be used to confirm a postmenopausal state in women not using hormonal contraception or hormonal replacement therapy (HRT). However, in the absence of 12 months of amenorrhea, a single follicle stimulating hormone measurement is insufficient.
 - Females on HRT and whose menopausal status is in doubt will be required to use one of the non-estrogen hormonal highly effective contraception methods if they wish to continue their HRT during the study. Otherwise, they must discontinue HRT to allow confirmation of postmenopausal status before study enrollment.

Contraception Guidance

Female participants of childbearing potential are eligible to participate if they agree to use a highly effective method of contraception consistently and correctly as described in the table below.

Highly Effective Contraceptive Methods^a

Highly Effective Contraceptive Methods That Are User Dependent^b

Failure rate of <1% per year when used consistently and correctly.

Combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation

- Oral
- Intravaginal
- Transdermal

Progestogen only hormonal contraception associated with inhibition of ovulation

- Oral
- Injectable

Highly Effective Methods That Are User Independent

Implantable progestogen only hormonal contraception associated with inhibition of ovulation

- Intrauterine device
- Intrauterine hormone-releasing system
- Bilateral tubal occlusion

Vasectomized partner

A vasectomized partner is a highly effective contraception method provided that the partner is the sole male sexual partner of the WOCBP and the absence of sperm has been confirmed. If not, an additional highly effective method of contraception should be used.

Sexual abstinence

Sexual abstinence is considered a highly effective method only if defined as refraining from heterosexual intercourse during the entire period of risk associated with the IMP. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study and the preferred and usual lifestyle of the participant.

NOTES:

- a) In case of newly started contraception pills/intrauterine devices, Investigator should consider the correct timing of starting/applying such methods in relation to the menstrual cycle and the manufacturing instruction as when these newly started methods would become effective.
- b) Typical use failure rates may differ from those when used consistently and correctly. Use should be consistent with local regulations regarding the use of contraceptive methods for participants participating in clinical studies.

Pregnancy testing

- WOCBP should only be included after a confirmed menstrual period and a negative highly sensitive serum pregnancy test.
- Additional pregnancy testing should be performed as indicated in the Schedule of Activities (Table 1–1) during the treatment period and at 20 weeks after the last dose of IMP and as required locally.
- Pregnancy testing will be performed whenever a menstrual cycle is missed or when pregnancy is otherwise suspected

Male participants with partners who become pregnant

- The Investigator will attempt to collect pregnancy information on any male participant's female partner who becomes pregnant while the male participant is in this study. This applies only to male participants who receive IMP. If the study participant is later found to be on placebo, then pregnancy data collection can stop.
- In cases where the partner of a male subject enrolled in a clinical study becomes pregnant, the Investigator or designee is asked to contact the subject to request consent of the partner via the Partner Pregnancy Consent form that has been approved by the responsible IRB/IEC and should be available in the Investigator site file. In case of questions about the consent process, the Investigator may contact the UCB/CRO contract monitor for the study. The Investigator will complete the information in the eCRF only after the partner has agreed that additional information can be captured and has provided the signed Partner Pregnancy Consent form. UCB's Patient Safety department is also the primary contact for any questions related to the data collection for the partner pregnancy, eventual birth, and follow up.
- After obtaining the necessary signed informed consent from the pregnant female partner directly, the Investigator will record pregnancy information on the appropriate form and submit it to the Sponsor within 1 working day of learning of the partner's pregnancy. The female partner will also be followed to determine the outcome of the pregnancy. Information on the status of the mother and child will be forwarded to the Sponsor. Generally, the follow-up will be at least 12 months after the delivery date. Any termination of the pregnancy will be reported regardless of fetal status (presence or absence of anomalies) or indication for the procedure.

Female Participants who become pregnant

- Any female participant who becomes pregnant while participating in the study will discontinue IMP or be withdrawn from the study.
- The Investigator will collect pregnancy information on any female participant who becomes pregnant while participating in this study. Information will be recorded on the appropriate form and submitted to the Sponsor within 1 working day of learning of a participant's pregnancy. The participant will be followed to determine the outcome of the pregnancy. The Investigator will collect follow-up information on the participant and the neonate and the information will be forwarded to the Sponsor. Generally, the follow-up will be at least 12 months after the delivery date. Any termination of pregnancy will be reported, regardless of fetal status (presence or absence of anomalies) or indication for the procedure.
- While pregnancy itself is not considered to be an AE or SAE, any pregnancy complication or elective termination of a pregnancy will be reported as an AE or SAE. A spontaneous abortion is always considered to be an SAE and will be reported as such. Any post-study pregnancy related SAE considered reasonably related to the IMP by the Investigator will be reported to the Sponsor as described in Section 8.3.5. While the Investigator is not obligated to actively seek this information in former study participants, he or she may learn of an SAE through spontaneous reporting.

10.5 Appendix 5: Genetics

Use and Analysis of DNA

Samples for potential future exploratory biomarker research will be collected and stored from consenting participants in the study. This sampling is optional for study participants and requires a separate informed consent. A decision not to consent does not exclude the study participant from the study:

- Blood sample for DNA.

These samples will only be used to further understanding of HS and/or how biomarkers, including genetic variation, may affect response or be affected by treatment with bimekizumab, background products, and/or concomitant medications in study participants with HS.

10.6 Appendix 6: Liver safety – suggested actions and follow-up assessments

Participants with potential drug-induced liver injury must be assessed to determine if IMP must be discontinued, as outlined in Section 7.1.2.

All PDILI events must be reported as an AE, and PDILI events meeting SAE criteria should be reported to the Sponsor within 24 hours of learning of the occurrence. Any PDILI event that meets the criterion for potential Hy's Law must be reported within 24 hours of learning of the occurrence as an AE of special interest (Section 8.3.6), and, if applicable, also reported as an SAE (Section 8.3).

Evaluation of PDILI consists of the diagnostic testing and continued monitoring included in Table 10–1 (specific tests dependent on laboratory results and corresponding symptoms) and consultation with a local hepatologist (if applicable; discussed in Section 10.6.1). The local hepatologist is the expert usually consulted by the treating physician for assessment and management of potential hepatic disease. This would usually be a hepatologist, but may be a gastroenterologist. Additional investigation and monitoring may be required and adapted based on the diagnosis after the cause of the liver injury/abnormality is confirmed (details in Section 7.1.2.1.1).

Table 10–1: Required investigations and follow-up for PDILI

Laboratory value			Immediate		Follow up	
ALT or AST	Total bilirubin	Symptoms ^a of hepatitis or hypersensitivity	Consultation requirements	Actions	Testing	Evaluation
≥3×ULN	≥2×ULN ^b	NA	Hepatology consult ^c Medical Monitor must be notified within 24 hours (eg, by laboratory alert) and subject discussed with Medical Monitor ASAP.	Immediate IMP discontinuation. ^d	Essential: Must have repeat liver chemistry values and additional testing completed ASAP (see Section 10.6.3); recommended to occur at the site with HCP.	Monitoring of liver chemistry values at least twice per week until values normalize, stabilize, or return to within baseline values. ^e
≥3×ULN	NA	Yes				
≥8×ULN	NA	NA				

Table 10–1: Required investigations and follow-up for PDILI

Laboratory value			Immediate		Follow up	
ALT or AST	Total bilirubin	Symptoms ^a of hepatitis or hypersensitivity	Consultation requirements	Actions	Testing	Evaluation
≥5×ULN (and ≥2× Baseline) and <8×ULN	<2×ULN	No	Discussion with Medical Monitor required. Consider need for hepatology consult if there is no evidence of resolution (see Follow-up requirements). ^c	<p>Further investigation – immediate IMP discontinuation not required (see Section 10.6.2).</p> <p>IMP discontinuation required if any of the following occur:</p> <ul style="list-style-type: none"> • Subject cannot comply with monitoring schedule. • Liver chemistry values continue to increase • Liver chemistry values remain ≥5×ULN (and ≥2× baseline) after 4 weeks of monitoring without evidence of resolution 	<p>Essential: Every attempt must be made to have repeat liver chemistry values and additional testing completed within 48 hours at the site with HCP (see Section 10.6.3).</p>	<p>Monitoring of liver chemistry values at least twice per week for 2 weeks.^e</p> <ul style="list-style-type: none"> • Immediate IMP discontinuation required if liver chemistry values continue to increase. <p>After 2 weeks of monitoring liver chemistry values:</p> <ul style="list-style-type: none"> • ALT or AST remains ≥5×ULN <8×ULN, IMP should be temporarily withheld and subject should undergo repeat test in 2 weeks. <p>Continue IMP if ALT or AST values <5×ULN; continue to monitor at least twice per week until values normalize, stabilize, or return to within baseline values.</p> <p>If ALT or AST remains ≥5×ULN after second retest, immediate IMP discontinuation required.</p> <p>Continue to monitor until values normalize, stabilize, or return to within baseline values.^d</p>

Table 10–1: Required investigations and follow-up for PDILI

Laboratory value			Immediate		Follow up	
ALT or AST	Total bilirubin	Symptoms ^a of hepatitis or hypersensitivity	Consultation requirements	Actions	Testing	Evaluation

ALP=alkaline phosphatase; ALT=alanine aminotransferase; ASAP=as soon as possible; AST=aspartate aminotransferase; HCP=healthcare practitioner; IMP=investigational medicinal product; NA=not applicable; PDILI=potential drug induced liver injury; ULN=upper limit of normal

^a Hepatitis symptoms include fatigue, nausea, vomiting, and right upper quadrant pain or tenderness; hypersensitivity symptoms include eosinophilia (>5%), rash, and fever (without clear alternative cause).

^b If the subject also has $\geq 2 \times$ ULN ALP, the possibility of an indication of biliary obstruction should be discussed with the Medical Monitor.

^c Details provided in Section 10.6.1 . The local hepatologist is the expert usually consulted by the treating physician for assessment and management of potential hepatic disease. This would usually be a hepatologist, but may be a gastroenterologist.

^d Details are provided in Section 10.6.2.

^e Unless an alternative monitoring schedule is agreed by the Investigator and UCB responsible physician, Determination of stabilization is at the discretion of the Investigator in consultation with the hepatologist (as applicable) and UCB responsible physician, as needed.

The results of all monitoring, including laboratory testing and other testing, should be made available to the study site and Sponsor.

All initial tests resulting in abnormal hepatic laboratory values need to be repeated, but appropriate medical action must not be delayed waiting for the repeat result.

If tests are done locally for more rapid results, a concurrent sample should also be sent to the central laboratory whenever possible. Medical care decisions are to be made initially using the most rapidly available results and a conservative approach must be taken if the results from the 2 laboratory tests are significantly different. Data from the local and central laboratory are to be recorded on the applicable eCRF pages.

When IMP is discontinued, all concomitant medications and herbal supplements that are not medically necessary should also be discontinued. In these cases, the Investigator should also consider dose reduction for medically necessary concomitant medication and consider changing any medically required concomitant medication known to be hepatotoxic to a suitable alternative.

When IMP is stopped due to PDILI (as described in Section 7.1.2.1), IMP must be permanently discontinued unless a subsequent alternative diagnosis fully explains the hepatic findings. If a subsequent alternative diagnosis fully explains the hepatic findings, and the requirements provided in Section 10.6.2 are met, rechallenge with IMP may be appropriate.

The approach to investigate PDILI is summarized in Table 10-1.

10.6.1 Consultation with Medical Monitor and local hepatologist

Potential drug-induced liver injury events require notification of the Medical Monitor or UCB study physician within 24 hours (eg, by laboratory alert), and the study participant must be discussed with the Medical Monitor or UCB study physician as soon as possible. If required, the study participant must also be discussed with the local hepatologist. The local hepatologist is the expert usually consulted by the treating physician for assessment and management of potential hepatic disease. This would usually be a hepatologist, but may be a gastroenterologist. If determined necessary, this discussion should be followed by a full hepatology assessment (see Section 10.6.3) and SAE report (if applicable).

10.6.2 Immediate action: determination of IMP discontinuation

All PDILI events require immediate action, testing, and monitoring.

The immediate action is dependent on the laboratory values and symptoms of hepatitis or hypersensitivity and ranges from continuation of IMP (followed by immediate investigation) to immediate and permanent discontinuation (see Section 7.1.2 and Table 10-1 for details).

When IMP is discontinued, all concomitant medications and herbal supplements that are not medically necessary should also be discontinued. The Investigator should also consider dose reduction for medically necessary concomitant medication and consider changing any medically required concomitant medication known to be hepatotoxic to a suitable alternative.

10.6.2.1 IMP restart/rechallenge

Study participants who are immediately discontinued from IMP due to having met certain criteria for PDILI (as described in Section 7.1.2 and Table 10-1), but for whom an alternative

diagnosis is confirmed, ie, drug-induced liver injury is excluded, can rarely restart IMP. Rechallenge with IMP can occur only if ALL of the following requirements are met:

- The results of additional testing and monitoring described in Section 10.6.3 and Section 7.1.2.1.1 confirm a nondrug-related cause for the abnormal hepatic laboratory parameters and any associated symptoms (ie, a subsequent alternative diagnosis fully explains the hepatic findings).
- No alternative treatment options are available to the study participant.
- The study participant has shown clear therapeutic benefit from the IMP.
- Study participant's ALT or AST elevations do not exceed $\geq 3 \times \text{ULN}$.
- Study participant's total bilirubin is $< 1.5 \times \text{ULN}$.
- Study participant has no signs or symptoms of hypersensitivity.
- The rechallenge is approved by the UCB study physician and a hepatologist. The hepatologist must be external to UCB. It is recommended that the hepatologist be a local hepatology expert or the hepatologist treating the study participant.
- Study participant agrees to the Investigator-recommended monitoring plan and understands his or her individual benefit risk for restarting IMP and this is adequately documented.

10.6.3 Testing: identification/exclusion of alternative etiology

The measurements and additional information required for the assessment of PDILI events when there is a reasonable possibility that they may have been caused by the IMP are detailed in Table 10-2 (laboratory measurements) and Table 10-3 (additional information). Results of the laboratory measurements and information collected are to be submitted to the Sponsor on the corresponding eCRF. If the medical history of the study participant indicates a requirement for other assessments not included below, these additional assessments should be completed and submitted, as applicable.

All blood samples should be stored, if possible. If tests are done locally for more rapid results, a concurrent sample must also be sent to the central laboratory.

Table 10–2: PDILI laboratory measurements

Virology-related	Hepatitis A IgM antibody
	HBsAg
	Hepatitis E IgM antibody
	HBcAb-IgM
	Hepatitis C RNA
	Cytomegalovirus IgM antibody
	Epstein-Barr viral capsid antigen IgM antibody (if unavailable, obtain heterophile antibody or monospot testing)
Immunology	Anti-nuclear antibody (qualitative and quantitative)
	Anti-smooth muscle antibody (qualitative and quantitative)
	Type 1 anti-liver kidney microsomal antibodies (qualitative and quantitative)
Hematology	Eosinophil count
Urinalysis	Urine drug screen ^a
Chemistry	Amylase
	Sodium, potassium, chloride, glucose, blood urea nitrogen, creatinine
	Total bilirubin, ALP, AST, ALT, gamma-glutamyltransferase, total cholesterol, albumin
	If total bilirubin $\geq 1.5 \times \text{ULN}$, obtain fractionated bilirubin to obtain % direct bilirubin
	Serum creatine phosphokinase and lactate dehydrogenase to evaluate possible muscle injury causing transaminase elevation
Additional	Prothrombin time/INR ^b
	Serum pregnancy test ^c
	PK sample

ALP=alkaline phosphatase; ALT=alanine aminotransferase; AST=aspartate aminotransferase; HBcAb-IgM=hepatitis B core antibody-IgM; HBsAg=hepatitis B surface antigen; IgM=immunoglobulin M; INR=international normalized ratio; PDILI=potential drug-induced liver injury; PK=pharmacokinetic; RNA=ribonucleic acid; ULN=upper limit of normal

^a Tests in addition to the specified analytes may be performed based on the Investigator's medical judgment and study participant history.

^b Measured only for study participants with ALT $> 8 \times \text{ULN}$, elevations in total bilirubin, and symptoms of hepatitis or hypersensitivity. Hepatitis symptoms include fatigue, nausea, vomiting, and right upper quadrant pain or tenderness; hypersensitivity symptoms include eosinophilia ($> 5\%$), rash, and fever (without clear alternative cause).

^c For women of childbearing potential.

Additional information to be collected is presented in [Table 10–3](#).

Table 10–3: PDILI information to be collected

New or updated information
<ul style="list-style-type: none"> Concomitant prescription and over-the-counter medications (eg, acetaminophen, herbal remedies, vitamins); dosages and dates should be included.
<ul style="list-style-type: none"> Pertinent medical history, including the following: <ul style="list-style-type: none"> History of liver disease (eg, autoimmune hepatitis, nonalcoholic steatohepatitis or other “fatty liver disease”) Adverse reactions to drugs Allergies Relevant family history or inheritable disorders (eg, Gilbert’s syndrome, alpha-1 antitrypsin deficiency) Recent travel Progression of malignancy involving the liver (Note: Metastatic disease to the liver, by itself, should not be used as an explanation for significant AST and/or ALT elevations.)
<ul style="list-style-type: none"> The appearance or worsening of clinical symptoms of hepatitis or hypersensitivity (eg, fatigue, nausea, vomiting, right upper quadrant pain or tenderness, decreased appetite, abdominal pain, jaundice, fever, or rash)
<ul style="list-style-type: none"> Recent clinically significant hypotension or hypoxemia with compromised cardiopulmonary function
<ul style="list-style-type: none"> Alcohol and illicit drug use
<ul style="list-style-type: none"> Results of liver imaging or liver biopsy, if done
<ul style="list-style-type: none"> Results of any specialist or hepatology consult, if done
<ul style="list-style-type: none"> Any postmortem/pathology reports

ALT=alanine aminotransferase; AST=aspartate aminotransferase; PDILI=potential drug-induced liver injury

**10.7 Appendix 7: Medical device AEs, Adverse device effects, SAEs,
and device deficiencies: definition and procedures for
recording, evaluating, follow-up, and reporting**

Not applicable to this study.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.8 Appendix 8: Rapid alert procedures

Not applicable to this study.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.9 Appendix 9: Country-specific requirements

Country-specific requirements will be provided separately, as applicable.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.10 Appendix 10: Abbreviations and trademarks

ADAb	anti-drug antibody
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AN	abscess and inflammatory nodule
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
CAT	computed axial tomography
COVID-19	coronavirus disease 2019
CPK	creatine phosphokinase
CRF	case report form
CRO	contract research organization
CTCAE	Common Terminology Criteria for Adverse Events
CXR	chest x-ray
DLQI	Dermatology Life Quality Index
DMC	Data Monitoring Committee
DNA	deoxyribonucleic acid
ECG	electrocardiogram
eCRF	electronic Case Report form
eC-SSRS	electronic Columbia-Suicide Severity Rating Scale
EQ-5D-3L	European Quality-of-Life 5 dimensions-3 level questionnaire
EudraCT	European Union Drug Regulating Authorities Clinical Trials
FAS	Full Analysis Set
GCP	Good Clinical Practice
HCV	hepatitis C virus
HiSCR	Hidradenitis Suppurativa Clinical Response
HiSCR ₂₅	a 25% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSCR ₅₀	a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSCR ₇₅	a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count

HiSCR ₉₀	a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSCR ₁₀₀	a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSQOL	Hidradenitis Suppurativa Quality of Life
HRT	hormonal replacement therapy
HS	hidradenitis suppurativa
hs-CRP	high-sensitivity C-reactive protein
HSSDD	Hidradenitis Suppurativa Symptom Daily Diary
HSSQ	Hidradenitis Suppurativa Symptom Questionnaire
IBD	inflammatory bowel disease
ICF	informed consent form
ICH	International Council for Harmonisation
IEC	Independent Ethics Committee
Ig	immunoglobulin
IGRA	interferon gamma release assay
IHS4	International Hidradenitis Suppurativa Severity score system
IL	interleukin
IMP	investigational medicinal product
IND	Investigational New Drug
IRB	Institutional Review Board
IRT	interactive response technology
LTBI	latent tuberculosis infection
<i>M.</i>	<i>Mycobacterium</i>
mAb	monoclonal antibody
NTM	nontuberculous mycobacterial
NRS	numeric rating scale
NSAID	nonsteroidal anti-inflammatory drug
PDILI	potential drug-induced liver injury
PEOT	Premature End of Treatment Visit
PGI-C-HS	Patient Global Impression of Change in Hidradenitis Suppurativa Severity
PGI-C-SP	Patient Global Impression of Change in Severity of Skin Pain

PGI-S-HS	Patient Global Impression of Hidradenitis Suppurativa Severity
PGI-S-SP	Patient Global Impression of Severity of Skin Pain
PHQ-9	Patient Health Questionnaire Depression Module-9
PK	pharmacokinetic(s)
PRN	as needed
PRO	patient-reported outcome
PSO	psoriasis
Q2W	every 2 weeks
Q4W	every 4 weeks
QOL	quality of life
RS	Randomized Set
SAE	serious adverse event
SAP	statistical analysis plan
sc	subcutaneous
SFU	Safety Follow-up
TEAE	treatment-emergent adverse event
TB	tuberculosis
TSQM-9	Treatment Satisfaction Questionnaire – Medication 9
ULN	upper limit of normal
WOCBP	woman of childbearing potential
WPAI-SHP	Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem

10.11 Appendix 11: Protocol amendment history

The Protocol Amendment Summary of Changes Table for the current amendment (Amendment 5) is located directly before the Table of Contents.

Amendment 4 (09 May 2022)

Overall rationale for the amendment

It is recommended that a threshold for within-patient clinically meaningful change to define treatment success be used in order to establish efficacy for skin pain in Phase 3 trials of patients with moderate to severe hidradenitis suppurativa. The Sponsor conducted analyses to estimate the threshold for within-patient clinically meaningful change that can be used for a responder definition based on the Hidradenitis Suppurativa Symptom Daily Diary worst skin pain item score using established guidelines and analytical methods. Pain response status at Week 16 using this definition has been added as a secondary endpoint to the study.

Section # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3 Objectives and endpoints	<ul style="list-style-type: none"> An additional secondary efficacy endpoint was added to the protocol assessing pain response at Week 16 using a threshold for within-patient clinically meaningful change as recommended by the U.S. Food and Drug Administration. Other pain endpoints clarified and updated accordingly. 	To align with FDA recommendations.
9.3.2 Analysis of the secondary efficacy endpoints	<ul style="list-style-type: none"> Section updated to include the new pain response endpoint in the statistical hierarchy, and description of planned analysis Figure 9-1 updated to include the new pain response endpoint. 	To accommodate the new pain response endpoint.
9.8 Determination of sample size	<ul style="list-style-type: none"> Sample size assumptions updated based on additional secondary endpoint. 	The addition to the protocol of the HSSDD worst skin pain response as a ranked secondary endpoint resulted in an addition to the power calculation section, including assumptions for sample size, response rates, and statistical power.
Throughout	<ul style="list-style-type: none"> Minor editorial and formatting revisions 	Minor edits and formatting revisions that do not impact content were made for readability and/or clarity

Amendment 3 (03 Feb 2021)

Overall Rationale for the Amendment

The main reason for this protocol amendment is due to Regulatory Agency feedback and to provide procedural clarifications.

Section # and Name	Description of Change	Brief Rationale
Title page	<ul style="list-style-type: none"> National Clinical Trial (NCT) number has been added 	NCT number was received on 24 Jan 2020
Serious adverse event reporting	<ul style="list-style-type: none"> Fax and email for Japan have been removed 	Deleted as there are no study sites in Japan
1.1 Synopsis 2.1 Study rationale 2.2 Background	<ul style="list-style-type: none"> Updated description of bimekizumab and other minor edits Order of secondary efficacy endpoints aligned with closed testing procedure (Figure 9-1) 	Rationale is the same as the description
1.1 Synopsis (Treatment Groups and Duration)	<ul style="list-style-type: none"> Updated Baseline antibiotic therapy strata to remove the 30% cap on enrollment [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] 	<p>Removed the Baseline antibiotic strata cap per US FDA advice (Baseline antibiotic use should coincide with the prevalence of antibiotic use for the disease)</p> <p>To align with the final Data Monitoring Committee Charter (DMC) Charter and DMC Statistical Analysis Plan (SAP)</p>
1.3 Schedule of activities	<ul style="list-style-type: none"> Extended the line for concomitant medications and adverse events to show they are collected through the Safety Follow-Up Visit Added an additional footnote that past medical history includes tobacco and alcohol use Added Note in footnotes that study assessments could be completed remotely in exceptional circumstances Footnote “n” (formally footnote “m”) has been updated regarding collection of Hurley Stage 	<p>To align collection of adverse events and concomitant medications with the protocol body text</p> <p>To include/clarify that tobacco and alcohol use is part of the medical history</p> <p>Provided operational flexibility to allow assessments to be collected remotely due to COVID-19 or other exceptional circumstance (eg, hurricanes)</p>

Section # and Name	Description of Change	Brief Rationale
		Clarification to make footnote "n" consistent with the table regarding collection of Hurley Stage
3 Objectives and endpoints Other	<ul style="list-style-type: none"> Order of secondary efficacy endpoints aligned with closed testing procedure (Figure 9-1) Other safety topics of interest endpoints were updated 	<p>With progressive development of bimekizumab and based on the ongoing review of emerging safety data, depression has been removed as a safety topic of interest. Nevertheless, depression will continue to be monitored as a safety parameter by the Patient Health Questionnaire 9 (PHQ-9) and will be captured via routine adverse event (AE) reporting during the study. This update is considered a procedural change.</p> <p>Rational for reordering the secondary efficacy endpoints is the same as the description.</p>
4.1 Overall design 4.2 Scientific rationale for study design	<ul style="list-style-type: none"> Updated Baseline antibiotic strata to remove the 30% cap on enrollment 	Removed the Baseline antibiotic strata cap per US FDA advice (Baseline antibiotic use should coincide with the prevalence of antibiotic use for the disease)
5.1 Inclusion criteria	<ul style="list-style-type: none"> Criteria #2 added text that diagnosis must be verifiable through medical notes and documentation Criteria #5 was edited for clarity, and added text that diagnosis/inadequate response must be verifiable through medical notes and documentation 	Clarification of criteria
5.2 Exclusion criteria	<ul style="list-style-type: none"> Criterion #16 text was updated to clarify use of the Screening Version of the electronic Columbia-Suicidality Severity Rating Scale (eC-SSRS) 	Clarification of criterion #16 as it is actually collected and assessed

Section # and Name	Description of Change	Brief Rationale
	<ul style="list-style-type: none"> Criterion #19 was updated to exclude participants with prior use of an IL-17 biologic response modifier or participation in an IL-17 biologic response modifier study unless an appropriate washout period (within 6 months prior to the Baseline Visit or 5 half-lives, whichever is greater) has been performed 	Criterion #19 was modified to allow enrollment of a moderate to severe HS population with real-world prior use of other medications with appropriate washout periods
5.4 Screen failures	<ul style="list-style-type: none"> Added bullet explaining that study participants who require incision and drainage procedures for HS lesions are to be screen failed 	Updated the screen failure criteria
6.5.1 Permitted concomitant treatments	<ul style="list-style-type: none"> Wound care updated to add that use of wound care dressings will be recorded Added Lesion care Updated Baseline antibiotic strata to remove the 30% cap on enrollment 	<p>Clarifications of and additions to allowed concomitant treatments</p> <p>Removed the Baseline antibiotic strata cap per US FDA advice (Baseline antibiotic use should coincide with the prevalence of antibiotic use for the disease)</p>
6.5.2 Prohibited concomitant treatments	<ul style="list-style-type: none"> Added washout periods for systemic antibiotics if applicable; other biologics; IL-17, IL-12, and IL-23 inhibitors; and janus kinase inhibitors Filgotinib and Upadacitinib added under janus kinase inhibitors Text added to clarify topical drugs Added herbal medications for HS and a washout period Updated vaccine criteria Text added to clarify that medications listed are currently available medications, but the protocol will account for medicine approvals in a given class during the course of the study 	Clarification of prohibited medications/therapies and the criteria for their exclusion and to allow enrollment of a moderate to severe HS population with real-world prior use of other medications (current and future) with appropriate washout periods
7.1 Discontinuation of study medication	<ul style="list-style-type: none"> Added a paragraph to clarify procedures if a study participant tested positive for COVID-19 or a suspected COVID-19 infection 	Updated for the COVID-19 pandemic
7.1.2.1 PDILI discontinuation criteria	<ul style="list-style-type: none"> Removed 'and permanent' from the first sentence and a crossreference to Section 10.6.2.1 has been included 	To be consistent with the PDILI criteria throughout the protocol and Appendix 10.6

Section # and Name	Description of Change	Brief Rationale
7.1.3 Treatment interruptions	<ul style="list-style-type: none"> Text was updated to clarify that doses of IMP that were missed due to a reasonable interfering AE that does not allow administration of IMP due to safety reasons, will not be considered for the evaluation of study participant discontinuation. It is still used to calculate compliance. A specific infection-related IMP interruption criterion has been added 	<p>Clarification of study procedures.</p> <p>In line with the exclusion criterion regarding infections, a specific infection-related IMP interruption criterion was added to clarify that participants with serious or recurrent infections not responding to standard therapies are not exposed to immunomodulatory therapies until their infection has resolved. This is in line with most biologic therapies, including other anti-IL17s.</p>
7.2 Participant discontinuation/withdrawal from the study	<ul style="list-style-type: none"> Added cross-reference to pregnancy section 	Clarification of study procedures
8 Study Assessments and Procedures	<ul style="list-style-type: none"> Added text about allowing study assessments to be performed remotely during a pandemic or other exceptional circumstance 	Updated to allow study to proceed during COVID-19, and other exceptional circumstances (eg, hurricanes)
8.1.1 Lesion count	<ul style="list-style-type: none"> Removed “hypertrophic” from description of scars and added “HS lesions” 	Clarification of lesion definition
8.1.4.6 Euro-Quality of Life 5-Dimensions, 3 levels	<ul style="list-style-type: none"> Removed sentences #2 and #3 regarding summary index scores Added a clarification to the last sentence 	Sentence removed in line with SAP update as index scores are not required for the clinical study report
8.2.2 Vital signs	<ul style="list-style-type: none"> Noncontact forehead added to body temperature measurement 	Updated for the COVID-19 pandemic and to align with the electronic Case Report Form (eCRF)
8.2.6.1 Assessments at Screening	<ul style="list-style-type: none"> Specified that the TB questionnaire is administered by the Investigator or their designee 	Clarification of study procedure

Section # and Name	Description of Change	Brief Rationale
8.2.6.3.3 Tuberculosis questionnaire	<ul style="list-style-type: none"> Specified that the questionnaire is administered by the Investigator or their designee 	Clarification of study procedure
8.3 Adverse events and serious adverse events	<ul style="list-style-type: none"> Added statement that cases of COVID-19 infection will be recorded as AEs (or SAEs , as required) 	Updated for the COVID-19 pandemic
8.3.5 Pregnancy	<ul style="list-style-type: none"> Bulletpoint #2 “or be down-titrated as instructed at the early discontinuation visit” was deleted 	Down titration of bimekizumab is not required if IMP needs to be discontinued
8.3.7 Other safety topics of interest	<ul style="list-style-type: none"> Deleted depression from list of AEs considered safety topics of interest Updated “liver function test changes/enzyme elevations” to “hepatic events and potential drug-induced liver injury (PDILI)” Clarified major “adverse” cardiovascular events 	<p>With progressive development of bimekizumab and based on the ongoing review of emerging safety data, depression has been removed as a safety topic of interest. Nevertheless, depression will continue to be monitored as a safety parameter by the PHQ-9 and will be captured via routine AE reporting during the study. This update is considered a procedural change.</p> <p>To align with UCB internal documents regarding assessment of hepatic events and PDILI</p> <p>Typographical error corrected for major adverse cardiovascular events</p>
9.1.7 COVID-19 Free Set	<ul style="list-style-type: none"> Addition of subsection for a COVID-19 Free analysis set 	Updated for the COVID-19 pandemic
9.3.1 Analysis of the primary efficacy endpoint	<ul style="list-style-type: none"> Cross reference added to Section 9.3.4 for additional sensitivity analyses related to the assessment of COVID-19 pandemic on the primary efficacy endpoint analysis 	Updated for the COVID-19 pandemic

Section # and Name	Description of Change	Brief Rationale
9.3.2 Analysis of secondary efficacy endpoints	<ul style="list-style-type: none"> Points #3 and #4 have been reordered to be consistent with Figure 9-1 (Closed Testing Procedure) Missing figure caption was added 	Rational is the same as the description
9.3.4 COVID-19 impact analysis	<ul style="list-style-type: none"> Addition of subsection for COVID-19 impact analysis 	Updated for the COVID-19 pandemic
9.4.1 Safety analysis	<ul style="list-style-type: none"> Addition of sentence that Summaries of Confirmed and Suspected COVID-19 TEAEs, respectively, will be presented and their definitions will be provided in the SAP 	Updated for the COVID-19 pandemic
9.5 Handling of protocol deviations	<ul style="list-style-type: none"> Added statement that COVID-related protocol deviations would be listed and summarized separately 	Updated for the COVID-19 pandemic
9.7.1 Data Monitoring Committee	<ul style="list-style-type: none"> [REDACTED] 	To align with the final DMC Charter and DMC SAP
10.1.6 Data quality assurance	<ul style="list-style-type: none"> Added statement about performing some study-specific assessments remotely under certain exceptional circumstances 	Updated for the COVID-19 pandemic and other exceptional circumstances (eg, hurricanes)
10.4 Appendix 4: Contraceptive guidance and collection of pregnancy information	<ul style="list-style-type: none"> Deleted footnote "c" regarding hormonal contraception. "Hormonal contraception may be susceptible to interaction with the IMP, which may reduce the efficacy of the contraceptive method. In this case, 2 highly effective methods of contraception should be utilized during the treatment period and for at least 20 weeks after the last dose of IMP" 	To be consistent with other studies in the bimekizumab program
10.6 Appendix 6: Liver safety – suggested actions and follow-up assessments	<ul style="list-style-type: none"> Paragraph #2 has been updated to add that PDILI events meeting SAE criteria should be reported to the Sponsor within 24 hours of learning of the occurrence. The requirement to report to the study site has been removed. 	Correction of typographical error and clarification of reporting procedures
10.11 Appendix 11: Protocol amendment history	<ul style="list-style-type: none"> Updated with changes from previous global amendment 	Self-evident
Throughout	<ul style="list-style-type: none"> Minor editorial and formatting revisions 	Minor edits and formatting revisions that do not impact content

Section # and Name	Description of Change	Brief Rationale
		were made for readability and/or clarity

Amendment 2 (16 Dec 2019)

Overall Rationale for the Amendment

The main reason for global protocol amendment 2 was to update the study discontinuation/withdrawal criteria for study participants with IBD.

Section # and Name	Description of Change	Brief Rationale
7.2 Participant discontinuation/withdrawal from the study	<ul style="list-style-type: none"> Added text to IBD discontinuation/withdrawal criteria 	Previously approved text relating to discontinuation/withdrawal criteria for IBD was inadvertently removed from the original version of the protocol dated 29 Oct 2019. It is now being replaced
7.2 Participant discontinuation/withdrawal from the study	<ul style="list-style-type: none"> Fixed numbering in list of criteria 	Corrected typographical error in list numbering
10.11 Appendix 11: Protocol amendment history	<ul style="list-style-type: none"> Updated with changes from previous global protocol amendment 1 	Updated

Amendment 1 (06 Dec 2019)

Overall Rationale for the Amendment

The main reason for global protocol amendment 1 was to update the company name in line with the new Code of Companies and Associations recently adopted by Belgium.

Section # and Name	Description of Change	Brief Rationale
Title page	<ul style="list-style-type: none"> Updated company name from UCB Biopharma SPRL to UCB Biopharma SRL 	Change in company name on 02 Dec 2019
3 Objectives and Endpoints	<ul style="list-style-type: none"> Clarified wording of exploratory biomarker objective 	Clarification of objective with current genetics and biomarkers sections (Section 8.7 and Section 8.9, respectively)
8.2.6.3.7 Tuberculosis management of LTBI, active TB, or other NTB infection identified during study	<ul style="list-style-type: none"> Figure 8-1 was updated to reflect Screening terminology as follows: <ul style="list-style-type: none"> Green ovals that said “subject eligible for study continuation” or “subject eligible to continue the study” in original protocol were 	Figure 8-1 was corrected to reflect Screening terminology

Section # and Name	Description of Change	Brief Rationale
	<p>changed to say “subject eligible for study”</p> <ul style="list-style-type: none">– Red ovals that said “withdraw from the study” were changed to say “subject NOT eligible for study”	
10.11 Appendix 11: Protocol amendment history	<ul style="list-style-type: none">• Stated location of summary of changes table for the current amendment	Updated

11 REFERENCES

Alikhan A, Lynch PJ, Eisen DB. Hidradenitis suppurativa: a comprehensive review. *J Am Acad Dermatol*. 2009;60(4):539-61.

Basra MK, Chowdhury MM, Smith EV, Freemantle N, Piquet V. A review of the use of the dermatology life quality index as a criterion in clinical guidelines and health technology assessments in psoriasis and chronic hand eczema. *Dermatol Clin*. 2012;30(2):237-44.

Bharmal M, Payne K, Atkinson MJ, Desrosiers MP, Morisky DE, Gemmen E. Validation of an abbreviated Treatment Satisfaction Questionnaire for Medication (TSQM-9) among patients on antihypertensive medications. *Health Qual Life Outcomes*. 2009;7:36.

Caposiena C, Bianchi L. Flares in hidradenitis suppurativa in treatment with adalimumab. *J Eur Acad Dermatol Venereol* 2020;34:e138-9.

Cho KA, Suh JW, Lee KH, Kang JL, Woo SY. IL-17 and IL-22 enhance skin inflammation by stimulating the secretion of IL-1 β by keratinocytes via the ROS-NLRP3-caspase-1 pathway. *Int Immunol*. 2012;24(3):147-58.

Cosentyx [package insert]. East Hanover, NJ: Novartis Pharmaceuticals Corporation; 2019. <https://www.pharma.us.novartis.com/sites/www.pharma.us.novartis.com/files/cosentyx.pdf><https://www.pharma.us.novartis.com/sites/www.pharma.us.novartis.com/files/cosentyx.pdf>.

Danby FW, Margesson LJ. Hidradenitis suppurativa. *Dermatol Clin*. 2010;28(4):779-93.

Esmann S, Dufour DN, Jemec GB. Questionnaire-based diagnosis of hidradenitis suppurativa: specificity, sensitivity and positive predictive value of specific diagnostic questions. *Br J Dermatol*. 2010;163(1):102-6.

Finlay AY. Quality of life assessments in dermatology. *Semin Cutan Med Surg*. 1998;17(4):291-6.

Frew JW, Hawkes JE, Krueger JG. A systematic review and critical evaluation of immunohistochemical associations in hidradenitis suppurativa. *F1000Research*. 2018;7.

Humira [package insert]. North Chicago, IL: AbbVie Inc; 2019. <https://www.rxabbvie.com/pdf/humira.pdf>.

Hurley H. Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In: *Dermatologic surgery*. Roenigh R, Roenigh H, editors. Marcel Dekker. 1989:729-39.

Kimball AB, Alavi A, Jemec GBE, et al. Secukinumab in moderate to severe hidradenitis suppurativa: Primary endpoint analysis from the SUNSHINE and SUNRISE Phase 3 trials. Presented at: The European Academy of Dermatology and Venereology; September 2022; Milan, Italy.

Kimball AB, Kerdel F, Adams D, Mrowietz U, Gelfand JM, Gniadecki R. Adalimumab for the treatment of moderate to severe Hidradenitis suppurativa: a parallel randomized trial. *Ann Intern Med*. 2012;157(12):846-55.

Kimball AB, Jemec GB, Yang M, Kageliery A, Signorovitch JE, Okun MM, et al. Assessing the validity, responsiveness and meaningfulness of the Hidradenitis Suppurativa Clinical Response

(HiSCR) as the clinical endpoint for hidradenitis suppurativa treatment. *Br J Dermatol*. 2014;171(6):1434-42.

Kimball AB, Okun MM, Williams DA, Gottlieb AB, Papp KA, Zouboulis CC, et al. Two Phase 3 Trials of Adalimumab for Hidradenitis Suppurativa. *N Engl J Med*. 2016a;375(5):422-34.

Kimball AB, Sobell JM, Zouboulis CC, Gu Y, Williams DA, Sundaram M, et al. HiSCR (Hidradenitis Suppurativa Clinical Response): a novel clinical endpoint to evaluate therapeutic outcomes in patients with hidradenitis suppurativa from the placebo-controlled portion of a phase 2 adalimumab study. *J Eur Acad Dermatol Venereol*. 2016b;30(6):989-94.

Kimball AB, Okun MM, Williams DA, Gottlieb AB, Papp KA, Zouboulis CC, et al. Two Phase 3 Trials of Adalimumab for Hidradenitis Suppurativa. *N Engl J Med* 2016c;375:422-34.

Kirby JS, Moore B, Liephart P, Shumaker K, Mammis-Gierbolini A, Benhadou F, et al. A narrative review of the definition of 'flare' in hidradenitis suppurativa. *Br J Dermatol*. 2020; 182:24-8.

McMillan, K. Hidradenitis suppurativa: number of diagnosed patients, demographic characteristics, and treatment patterns in the United States. *Am J Epidemiol*. 2014;179(12):1477-83.

Martorell A, García-Martínez FJ, Jiménez-Gallo D, Pascual JC, Pereyra-Rodríguez J, Salgado L, Vilarrasa E. An update on hidradenitis suppurativa (part I): epidemiology, clinical aspects, and definition of disease severity. *Actas Dermo-Sifiligráficas (English Edition)*. 2015 Nov 1;106(9):703-15.

Mundt JC, Greist JH, Gelenberg AJ, Katzelnick DJ, Jefferson JW, Modell JG. Feasibility and validation of a computer-automated Columbia-Suicide Severity Rating Scale using interactive voice response technology. *J Psychiatr Res*. 2010;44(16):1224-8.

Naldi L. Epidemiology. In: *Hidradenitis Suppurativa*. Chapter 8. Jemec GB, Revuz J, Leyden J, editors. Springer-Verlag. 2006:58-64.

Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168(12):1266-77.

Prens E, Deckers I. Pathophysiology of hidradenitis suppurativa: an update. *J Am Acad Dermatol*. 2015;73(5):S8-11.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *PharmacoEconomics*. 1993;4(5):353-65.

Revuz JE, Canoui-Poitine F, Wolkenstein P, Vialette C, Gabison G, Pouget F, et al. Prevalence and factors associated with hidradenitis suppurativa: results from two case-control studies. *J Am Acad Dermatol*. 2008;59(4):596-601.

Riis PT, Søbey K, Saunte DM, Jemec GB. Patients with hidradenitis suppurativa carry a higher systemic inflammatory load than other dermatological patients. *Archives of dermatological research*. 2015 Dec 1;307(10):885-9.

Riis PT, Vinding GR, Ring HC, Jemec GB. Disutility in patients with hidradenitis suppurativa: a cross-sectional study using EuroQoL-5D. *Acta Derm Venereol*. 2016;96(2):222-6.

Sartorius K, Emtestam L, Jemec GB, Lapins J. Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *Br J Dermatol*. 2009;161(4):831-9.

Schlapbach C, Hänni T, Yawalkar N, Hunger RE. Expression of the IL-23/TH17 pathway in lesions of hidradenitis suppurativa. *J Am Acad Dermatol*. 2011;65(4):790-8.

Sun H, Snyder E, Koch GG. Statistical planning in confirmatory clinical trials with multiple treatment groups, multiple visits, and multiple endpoints. *J Biopharmaceut Stat*. 2018;28(1):189-211.

Thorlacius L, Ingram JR, Villumsen B, et al. A core domain set for hidradenitis suppurativa trial outcomes: an international Delphi process. *Br J Dermatol*. 2018.

Van der Zee HH, de Ruiter L, Van Den Broecke DG, Dik WA, Laman JD, Prens EP. Elevated levels of tumour necrosis factor (TNF)- α , interleukin (IL)-1 β and IL-10 in hidradenitis suppurativa skin: a rationale for targeting TNF- α and IL-1 β . *British Journal of Dermatology*. 2011 Jun;164(6):1292-8.

Vender R, Lynde C, Ho V, Chau D, Poulin-Costello M. Work productivity and healthcare resource utilization outcomes for patients on etanercept for moderate to severe plaque psoriasis: Results from a 1-year, multicentre, open-label, single-arm study in a clinical setting. *Appl Health Econ Health Policy*. 2012;10(5):343-53.

von der Werth JM, Jemec GB. Morbidity in patients with hidradenitis suppurativa. *Br J Dermatol*. 2001;144:809-13.

Wolkenstein P, Loundou A, Barrau K, Auquier P, Revuz J; Quality of Life Group of the French Society of Dermatology. Quality of life impairment in hidradenitis suppurativa: a study of 61 cases. *J Am Acad Dermatol*. 2007;56:621-3.

Zouboulis CC, Desai N, Emtestam L, Hunger RE, Ioannides D, Juhász I, et al. European S1 guideline for the treatment of hidradenitis suppurativa/acne inversa. *J Eur Acad Dermatol Venereol*. 2015;29(4):619-44.

Zouboulis CC, Tzellos T, Kyrgidis A. Development and validation of IHS4, a novel dynamic scoring system to assess hidradenitis suppurativa/acne inversa severity. *Br J Dermatol*. 2017;177:1401-9.

SPONSOR DECLARATION

I confirm that I have carefully read and understand this protocol and agree to conduct this clinical study as outlined in this protocol and according to current Good Clinical Practice.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Approval Signatures

Name: hs0003-protocol-amendment-5

Version: 1. 0

Document Number: CLIN-000204482

Title: HS0003-Protocol Amendment 5

Approved Date: 27 Sep 2022

Document Approvals	
Approval Verdict: Approved	Name: [REDACTED] Capacity: Medical Date of Signature: 27-Sep-2022 12:25:43 GMT+0000
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 27-Sep-2022 12:47:31 GMT+0000
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 27-Sep-2022 19:49:16 GMT+0000

**A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-
CONTROLLED, MULTICENTER STUDY EVALUATING THE
EFFICACY AND SAFETY OF BIMEKIZUMAB IN STUDY
PARTICIPANTS WITH MODERATE TO SEVERE
HIDRADENITIS SUPPURATIVA**

PROTOCOL HS0004 AMENDMENT 4

PHASE 3

SHORT TITLE:

A Phase 3, Randomized, Double-blind, Placebo-Controlled, Multicenter Study Evaluating the Efficacy and Safety of Bimekizumab in Study Participants with Moderate to Severe Hidradenitis Suppurativa

Sponsor:

UCB Biopharma SRL

Allée de la Recherche 60

1070 Brussels

BELGIUM

Regulatory agency identifying numbers:

Eudra CT Number:	2019-002551-42
Investigational New Drug (IND) Number:	134700
NCT Number:	NCT04242498

Confidential Material

Confidential

This document is the property of UCB and may not – in full or in part – be passed on, reproduced, published, or otherwise used without the express permission of UCB.

PROTOCOL AMENDMENT SUMMARY OF CHANGES TABLE

Document History		
Document	Date of Issue	Type of Amendment
Amendment 4	06 May 2022	Substantial
Amendment 3	09 Feb 2021	Substantial
Amendment 2.1 (Japan)	19 Dec 2019	Not applicable
Amendment 2	16 Dec 2019	Nonsubstantial
Amendment 1.1 (Japan)	10 Dec 2019	Not applicable
Amendment 1	06 Dec 2019	Nonsubstantial
Original Protocol	29 Oct 2019	Not applicable

Amendment 4 (06 May 2022)

Overall rationale for the amendment

It is recommended that a threshold for within-patient clinically meaningful change to define treatment success be used in order to establish efficacy for skin pain in Phase 3 trials of patients with moderate to severe hidradenitis suppurativa. The Sponsor conducted analyses to estimate the threshold for within-patient clinically meaningful change that can be used for a responder definition based on the Hidradenitis Suppurativa Symptom Daily Diary worst skin pain item score using established guidelines and analytical methods. Pain response status at Week 16 using this definition has been added as a secondary endpoint to the study.

Section # and Name	Description of Change	Brief Rationale
1.1 Synopsis 3 Objectives and endpoints	<ul style="list-style-type: none"> An additional secondary efficacy endpoint was added to the protocol assessing pain response at Week 16 using a threshold for within-patient clinically meaningful change as recommended by the U.S. Food and Drug Administration. Other pain endpoints clarified and updated accordingly. 	To align with FDA recommendations.
9.3.2 Analysis of the secondary efficacy endpoints	<ul style="list-style-type: none"> Section updated to include the new pain response endpoint in the statistical hierarchy, and description of planned analysis Figure 9-1 updated to include the new pain response endpoint. 	To accommodate the new pain response endpoint.
9.8 Determination of sample size	<ul style="list-style-type: none"> Sample size assumptions updated based on additional secondary endpoint. 	The addition to the protocol of the HSSDD worst skin pain response as a ranked secondary endpoint resulted in an addition to the power calculation section, including assumptions for sample size, response rates, and statistical power.
Throughout	<ul style="list-style-type: none"> Minor editorial and formatting revisions 	Minor edits and formatting revisions that do not impact content were made for readability and/or clarity

SERIOUS ADVERSE EVENT REPORTING

Serious adverse event reporting (24h)	
Fax	Europe and Rest of the World: +32 2 386 24 21 US and Canada: +1 800 880 6949 or +1 866 890 3175
Email	Global: DS_ICT@ucb.com (for interventional clinical studies)

TABLE OF CONTENTS

1	PROTOCOL SUMMARY	10
1.1	Synopsis	10
1.2	Schema	13
1.3	Schedule of activities	14
2	INTRODUCTION	20
2.1	Study rationale	20
2.2	Background	20
2.3	Benefit/risk assessment	21
3	OBJECTIVES AND ENDPOINTS	23
4	STUDY DESIGN	26
4.1	Overall design	26
4.1.1	Screening Period (Weeks -5 to 0)	27
4.1.2	Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)	27
4.1.3	Safety Follow-up Visit	27
4.1.4	Visit Windows	27
4.2	Scientific rationale for study design	27
4.3	Justification for dose	28
4.4	End of study definition	29
5	STUDY POPULATION	29
5.1	Inclusion criteria	29
5.2	Exclusion criteria	30
5.3	Lifestyle restrictions	34
5.4	Screen failures	34
6	STUDY TREATMENTS	35
6.1	Treatments administered	35
6.2	Preparation, handling, storage, and accountability requirements	38
6.2.1	Drug accountability	38
6.3	Measures to minimize bias: randomization and blinding	38
6.3.1	Randomization and numbering of participants	38
6.3.2	Procedures for maintaining and breaking the treatment blind	39
6.3.2.1	Maintenance of study treatment blind	39
6.3.2.2	Breaking the treatment blind in an emergency situation	39
6.4	Treatment compliance	40
6.5	Concomitant medications/treatments	40
6.5.1	Permitted concomitant treatments (medications and therapies)	40
6.5.2	Prohibited concomitant treatments (medications and therapies)	41

6.5.3	Rescue medication	44
6.5.3.1	Antibiotic Rescue Medication/Antibiotics	44
6.5.3.2	Rescue Medications/Lesion Intervention	44
6.6	Dose modification	45
6.7	Criteria for study hold or dosing stoppage.....	45
6.8	Treatment after the end of the study	45
7	DISCONTINUATION OF STUDY MEDICATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL.....	46
7.1	Discontinuation of study medication	46
7.1.1	Study participant does not achieve partial response	46
7.1.2	Potential drug-induced liver injury IMP discontinuation criteria	46
7.1.2.1	PDILI Discontinuation Criteria	46
7.1.3	Treatment interruptions	47
7.2	Participant discontinuation/withdrawal from the study	48
7.3	Lost to follow up	50
8	STUDY ASSESSMENTS AND PROCEDURES	51
8.1	Efficacy assessments	52
8.1.1	Lesion count.....	52
8.1.1.1	Hidradenitis Suppurativa Clinical Response (HiSCR).....	52
8.1.1.2	Hidradenitis Suppurativa Physician's Global Assessment.....	53
8.1.1.3	International Hidradenitis Suppurativa Severity Score System (IHS4)	53
8.1.2	Partial response	53
8.1.3	High-sensitivity C-reactive protein (hs-CRP)	54
8.1.4	Patient-reported outcomes	54
8.1.4.1	HS symptom measures of skin pain, smell or odor, drainage or oozing from HS lesions, and itch	54
8.1.4.2	Patient Global Impression of HS Severity (PGI-S-HS) and Change in HS Severity (PGI-C-HS)	54
8.1.4.3	Patient Global Impression of Severity of Skin Pain (PGI-S-SP) and Change in Severity of Skin Pain (PGI-C-SP).....	54
8.1.4.4	Hidradenitis Suppurativa Quality of Life Questionnaire (HiSQOL)	55
8.1.4.5	Dermatology Life Quality Index (DLQI)	55
8.1.4.6	Euro-Quality of Life 5-Dimensions, 3 levels	55
8.1.4.7	Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem.....	55
8.1.4.8	Treatment Satisfaction Questionnaire for Medication	55
8.1.5	Hurley Stage	56
8.2	Safety assessments	56
8.2.1	Physical examination	56

8.2.2	Vital signs	56
8.2.3	Electrocardiograms	57
8.2.4	Clinical safety laboratory assessments	57
8.2.5	Depression and suicidal risk monitoring	57
8.2.5.1	PHQ-9.....	57
8.2.5.2	eC-SSRS	57
8.2.6	Assessment and management of tuberculosis and tuberculosis risk factors.....	58
8.2.6.1	Assessments at Screening.....	58
8.2.6.2	Definitions	58
8.2.6.3	Assessment and reporting of TB and TB risk factors during the study.....	59
8.3	Adverse events and serious adverse events	64
8.3.1	Time period and frequency for collecting AE and SAE information.....	64
8.3.2	Method of detecting AEs and SAEs	64
8.3.3	Follow-up of AEs and SAEs.....	64
8.3.4	Regulatory reporting requirements for SAEs	65
8.3.5	Pregnancy	65
8.3.6	Adverse events of special interest.....	65
8.3.7	Other safety topics of interest	66
8.3.8	Anticipated serious adverse events	66
8.3.9	Suspected transmission of an infectious agent.....	67
8.4	Safety signal detection	67
8.5	Treatment of overdose	68
8.6	Pharmacokinetics	68
8.7	Genetics.....	68
8.8	Pharmacodynamics	69
8.9	Biomarkers	69
8.9.1	Immunogenicity assessments.....	69
8.10	Medical resource utilization and health economics	70
8.11	Photography	70
9	STATISTICAL CONSIDERATIONS.....	70
9.1	Definition of analysis sets.....	70
9.1.1	Enrolled Set	70
9.1.2	Randomized Set	70
9.1.3	Safety Set	70
9.1.4	Full Analysis Set.....	70
9.1.5	Per-Protocol Set	70
9.1.6	Pharmacokinetics Per-Protocol Set.....	71
9.1.7	COVID-19 Free Set	71

9.2	General statistical considerations.....	71
9.3	Planned efficacy/outcome analyses	71
9.3.1	Analysis of the primary efficacy endpoint.....	71
9.3.2	Analysis of the secondary efficacy endpoints	72
9.3.3	Other efficacy/other outcome analyses.....	74
9.3.4	COVID-19 impact analysis.....	74
9.4	Planned safety and other analyses.....	75
9.4.1	Safety analyses.....	75
9.4.2	PK and ADAb analyses	75
9.5	Handling of protocol deviations.....	75
9.6	Handling of dropouts or missing data.....	76
9.7	Planned interim analysis and data monitoring	77
9.7.1	Data Monitoring Committee.....	77
9.7.2	Interim Analysis.....	77
9.8	Determination of sample size.....	77
10	SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS... 80	
10.1	Appendix 1: Regulatory, ethical, and study oversight considerations.....	80
10.1.1	Regulatory and ethical considerations	80
10.1.2	Financial disclosure	80
10.1.3	Informed consent process.....	81
10.1.4	Data protection.....	81
10.1.5	Committees structure.....	82
10.1.6	Data quality assurance	82
10.1.6.1	Case report form completion.....	83
10.1.7	Source documents.....	83
10.1.8	Study and Site Closure.....	83
10.1.9	Publication Policy	84
10.2	Appendix 2: Clinical laboratory tests	85
10.3	Appendix 3: Adverse events – definitions and procedures for recording, evaluating, follow-up, and reporting	87
10.4	Appendix 4: Contraceptive guidance and collection of pregnancy information	93
10.5	Appendix 5: Genetics.....	96
10.6	Appendix 6: Liver safety – suggested actions and follow-up assessments	97
10.6.1	Consultation with Medical Monitor and local hepatologist.....	101
10.6.2	Immediate action: determination of IMP discontinuation	101
10.6.2.1	IMP restart/rechallenge	101
10.6.3	Testing: identification/exclusion of alternative etiology	102

10.7	Appendix 7: Medical device AEs, Adverse device effects, SAEs, and device deficiencies: definition and procedures for recording, evaluating, follow-up, and reporting.....	105
10.8	Appendix 8: Rapid alert procedures	106
10.9	Appendix 9: Country-specific requirements.....	107
10.10	Appendix 10: Abbreviations and trademarks	108
10.11	Appendix 11: Protocol amendment history	111
11	REFERENCES	118
	SPONSOR DECLARATION	121

LIST OF TABLES

Table 1–1:	Schedule of activities	14
Table 6–1:	Study medications administered	36
Table 6–2:	Dosing scheme	37
Table 6–3:	Exclusions for prior medications	42
Table 8–1:	Anticipated SAEs for the Population of Participants with HS	67
Table 9–1:	Power calculation assumptions and methods.....	79
Table 10–1:	Required investigations and follow-up for PDILI	98
Table 10–2:	PDILI laboratory measurements	103
Table 10–3:	PDILI information to be collected	104

LIST OF FIGURES

Figure 1-1:	Study Schematic.....	13
Figure 8-1:	Decision tree for IGRA TB results at Screening	62
Figure 8-2:	Decision tree for IGRA TB results during a study	63
Figure 9-1:	Closed Testing Procedure	74

1 **PROTOCOL SUMMARY**

1.1 **Synopsis**

Protocol title:

A Phase 3, Randomized, Double-blind, Placebo-Controlled, Multicenter Study Evaluating the Efficacy and Safety of Bimekizumab in Study Participants with Moderate to Severe Hidradenitis Suppurativa

Short Title:

A Phase 3, Randomized, Double-blind, Placebo-Controlled, Multicenter Study Evaluating the Efficacy and Safety of Bimekizumab in Study Participants with Moderate to Severe Hidradenitis Suppurativa

Rationale:

UCB is investigating bimekizumab (a humanized, full-length immunoglobulin [Ig] G1 monoclonal antibody [mAb] with high affinity for human interleukin [IL] 17 [IL-17A and IL 17F]) for the treatment of hidradenitis suppurativa (HS). Antibodies targeting IL-17A are effective in treating patients with moderate to severe psoriasis (PSO) and other immuno-inflammatory conditions. Data from clinical studies with bimekizumab suggest that inhibition of both IL-17A and IL-17F could be beneficial to patients with such conditions, including HS. Based on results from a Phase 2a study in study participants with HS (HS0001); and a Phase 2b study in study participants with PSO (PS0010) and its 48 week extension study (PS0011), bimekizumab doses of 320mg every 2 weeks (Q2W) in HS and 320mg every 4 weeks (Q4W) in PSO appeared to have an acceptable safety profile, and achieved clinically meaningful efficacy in both primary and key secondary endpoints in their respective studies. Based on these data and considering the high unmet need for safe and effective therapies for HS, confirmatory studies are being conducted with bimekizumab for the treatment of moderate to severe HS.

Objectives and Endpoints

The primary and secondary objectives and associated endpoints are as follows:

Objectives	Endpoints
Primary	
Evaluate the efficacy of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₅₀ at Week 16
Secondary	
Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₇₅ at Week 16 Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16 Absolute change from Baseline in DLQI Total Score at Week 16 Absolute change from Baseline in the Worst HS Skin Pain score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HSSDD Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) at Week 16
Evaluate the safety of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> Treatment-emergent AEs Serious TEAEs TEAEs leading to withdrawal from study

AE=adverse event; AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS= hidradenitis suppurativa; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; TEAE=treatment-emergent adverse event

Overall Design

HS0004 is a Phase 3, randomized, double-blind, placebo-controlled, multicenter, pivotal study evaluating the efficacy and safety of bimekizumab in study participants with moderate to severe HS.

Number of Participants

Approximately 490 study participants will be randomly assigned to study treatment: 140 for bimekizumab 320mg Q2W/Q2W, 140 for bimekizumab 320mg Q2W/Q4W, 140 for bimekizumab 320mg Q4W/Q4W, and 70 for placebo/320mg Q2W. The Randomized Set (ie, all randomized study participants) is the primary analysis set for efficacy analyses.

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomized to study treatment (Section 5.4).

Treatment Groups and Duration

Study participants meeting the inclusion criteria who do not meet any exclusion criteria will complete a Screening Period of 14 days to up to 5 weeks; a double-blind, 48-week Treatment Period comprising a 16-week Initial Treatment Period and 32-week Maintenance Treatment Period; and a 20-week Safety Follow-up (SFU) Period following the final injection of investigational medicinal product (IMP) if study participants do not enter a subsequent extension study (HS0005) or withdraw prematurely from treatment. The total duration of study participation in HS0004 will be 68 to 71 weeks for those who complete HS0004 and do not participate in the extension study HS0005 and 50 to 53 weeks for those who participate in HS0005 and, thus, do not participate in the 20-week SFU Period.

Study participants will be randomized in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to 1 of 3 dose regimens of bimekizumab or placebo as shown in the schematic (Figure 1-1). All doses of IMP will be administered by subcutaneous (sc) injection. The primary efficacy variable at Week 16 is HiSCR₅₀ (a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count). Study visits will occur at Screening; Baseline (Week 0); Weeks 1, 2, 4, 6, 8, 10, 12, 14, and 16; every 2 weeks from Week 16 through Week 48 for assessments of efficacy, safety, and other measures of quality of life/health status/work productivity; and a SFU visit 20 weeks after the last dose of IMP for participants who do not enter the extension study.

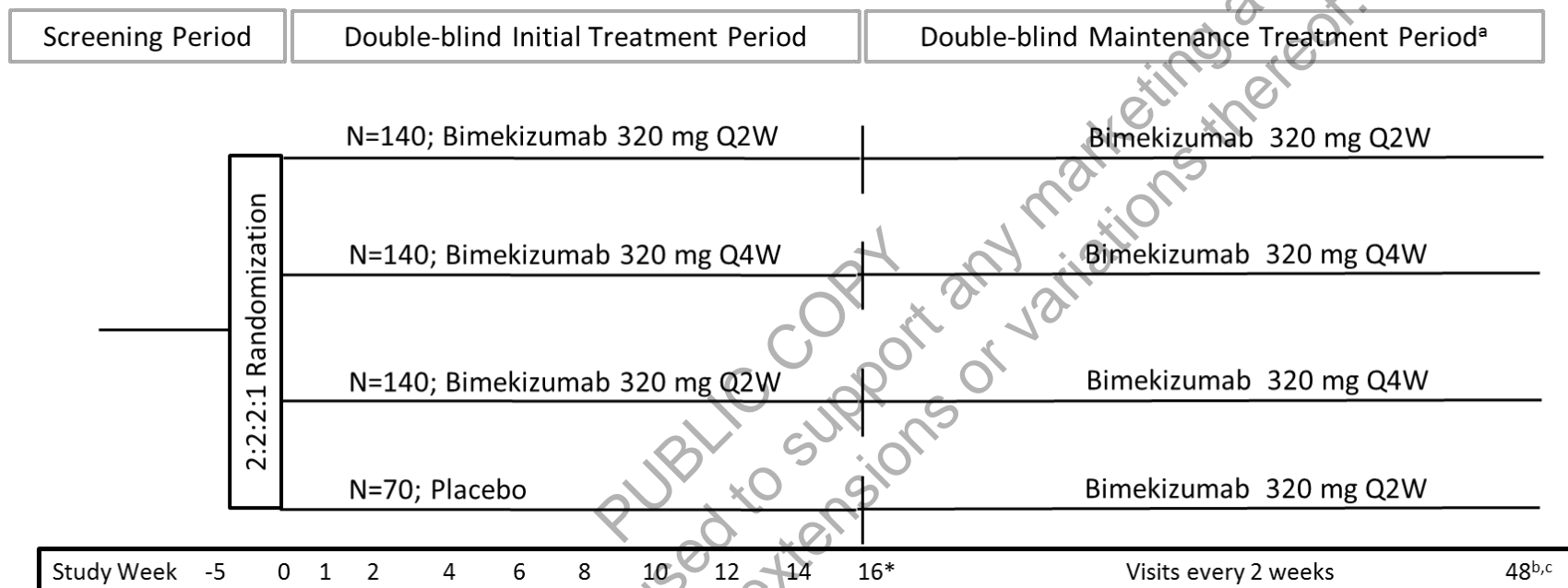
An independent Data Monitoring Committee (DMC) will periodically review unblinded efficacy and safety data to assess the benefit/risk of bimekizumab in study participants with moderate to severe HS. [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

1.2 Schema

A schematic diagram for the study is presented in [Figure 1-1](#).

Figure 1-1: Study Schematic



HiSCR₅₀=a 50% reduction in the total abscess and inflammatory nodule count with no increase from Baseline in abscess or draining tunnel count;

IMP=investigational medicinal product; Q2W=every 2 weeks; Q4W=every 4 weeks

*Week 16 = primary endpoint (HiSCR₅₀ bimekizumab versus placebo)

^a Study participants should discontinue from the study from Week 32 on if no partial response is achieved (partial response is defined as $\geq 25\%$ improvement in abscess and inflammatory nodule count relative to Baseline [Week 0] lesion values.)

^b Study participants achieving an improvement of at least 25 % in abscess and inflammatory nodule count continue in HS0005 (Extension Study).

^c 20-week Safety Follow-up (from last IMP injection) for any study participant who discontinues from study prior to Week 48, or who does not continue in HS0005.

1.3 Schedule of activities

The Schedule of Activities is presented in [Table 1–1](#).

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)											Maintenance Treatment Period (weeks after first dose)															PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Study Visit	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Informed consent	X																										X ^d		
Study participant number assigned	X																												
Inclusion/exclusion	X	X																											
Demographic and Baseline disease characteristics	X	X																											
Hidradenitis suppurativa history	X																												
Significant past medical history/concomitant diseases ^e	X	X ^f																											
Physical examination ^g	X	X							X		X				X						X						X	X	X

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFT ^c	
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48			
Height		X																												
Body weight		X												X													X	X		
Vital signs ^h	X	X	X	X	X		X		X		X	X	X	X		X		X		X		X		X		X	X	X		
12-lead ECG	X										X								X								X	X	X	
Hematology/ biochemistry	X	X		X	X		X		X		X	X	X	X		X		X		X		X		X		X	X	X		
hs-CRP		X							X		X								X							X	X			
Urinalysis	X	X									X			X					X				X			X	X	X		
Urine drug screen	X																									X	X			
Pregnancy testing ⁱ	X	X			X		X		X		X		X		X		X		X		X		X		X		X	X	X	
Hepatitis B and C testing ^j	X																													
HIV testing ^k	X																													
Chest x-ray ^l	X																													
IGRA TB test ^m	X																								X					
TB questionnaire	X	X							X					X						X							X	X	X	
Lesion count (includes Hurley Stage) ⁿ	X	X		X	X		X		X		X		X		X		X		X		X		X		X		X	X		

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Photography ^o		X			X						X								X								X	X	
Blood sample – bimekizumab plasma concentrations ^p		X	X	X	X		X		X		X	X	X		X						X						X	X	X
Blood sample for anti-drug antibodies ^p		X			X		X		X		X		X		X						X						X	X	X
Blood sample for genomic/ proteomic/ metabolomics, and candidate biomarker analyses ^{p, q}		X					X		X		X																X	X	
Blood sample for genetic/ epigenetic analyses ^{p, q}		X									X																X	X	
Urine samples for biomarker research		X									X																X	X	
DLQI		X			X		X		X		X		X						X		X						X	X	
PHQ-9	X	X			X		X		X		X		X		X		X		X		X		X		X		X	X	X

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
eC-SSRS	X	X			X		X		X		X		X		X		X		X		X		X		X		X	X	X
WPAI-SHP		X							X		X								X								X	X	
TSQM-9											X																X	X	
HiSQOL		X			X						X								X								X	X	
HSSDD ^r	●										●																		
HSSQ		X									X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PGI-S-HS		X			X						X								X								X	X	
PGI-C-HS					X						X								X								X	X	
PGI-S-SP		X			X						X								X								X	X	
PGI-C-SP					X						X								X								X	X	
EQ-5D-3L		X			X					X	X								X		X						X	X	
Concomitant medications	●																												
Adverse events	●																												
Randomization ^s		X																											
IRT	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Bimekizumab/ placebo administration ^t		X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			

AN=abscess and inflammatory nodule; Bsl=Baseline; CAT=computed axial tomography; CRF=case report form; DLQI=Dermatology Life Quality Index;

ECG=electrocardiogram; eC-SSRS=electronic Columbia-Suicide Severity Rating Scale; EQ-5D-3L=Euro-Quality of Life 5-Dimensions, 3 levels;

HiSCR₂₅=a 25% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₅₀=a 50% reduction in the total

Table 1–1: Schedule of activities

Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)																PEOT	SFU ^c
		Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46	48		
Protocol activity																													

AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₉₀=a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₁₀₀=a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSQOL=Hidradenitis Suppurativa Quality of Life; HIV=human immunodeficiency virus; HS=hidradenitis suppurativa; hs-CRP=high sensitivity C-reactive protein; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; HSSQ=Hidradenitis Suppurativa Symptom Questionnaire; IGRA=interferon-gamma release assay; IMP=investigational medicinal product; IRT=interactive response technology; PEOT=Premature End of Treatment Visit; PGI-C-HS=Patient Global Impression of Change in HS Severity; PGI-C-SP=Patient Global Impression of Change in Severity of Skin Pain; PGI-S-HS=Patient Global Impression of HS Severity; PGI-S-SP=Patient Global Impression of Severity of Skin Pain; PHQ-9=Patient Health Questionnaire 9; SFU=Safety Follow-Up; TB=tuberculosis; TSQM=Treatment Satisfaction Questionnaire for Medication; WPAI-SHP=Work Productivity and Impairment Questionnaire-Specific Health Problem

Note: Some study-specific assessments may be conducted remotely according to the study protocol during a pandemic or other exceptional circumstance (eg, hurricanes) due to the need to implement safety measures and guidance from regulatory authorities. Refer to Section 8 for additional information.

^a Visit windows are ± 3 days (based on the date of the first dose). However, the minimum of 8 days between injection visits (eg, Visit 2 +3 days occurs on Day17; Visit 4 -3 days occurs on Day 25) may be used only 1 time in the Initial Treatment Period and 1 time in the Maintenance Treatment Period, if needed. The study participant should be dosed according to the administration schedule thereafter. The 20-week SFU Visit window is ± 7 days (based on the date of the final dose).

^b The Screening Period will last a minimum of 14 days up to a maximum of 5 weeks prior to randomization.

^c All study participants not continuing in the extension study, including those withdrawn from IMP as well as those completing all study visits, will have an SFU Visit 20 weeks after his or her final dose of IMP.

^d A separate informed consent form will be required to be completed for participation in extension study HS0005. Study participants continuing in HS0005 will receive first dose of bimekizumab in that study on Week 48.

^e Past medical history includes tobacco and alcohol use.

^f Ensure no significant changes from Baseline in medical history and concomitant disease.

^g The physical examination will be performed as detailed in Section 8.2.1. Includes evaluation of signs and symptoms of active TB and risk for exposure to TB.

^h Vital signs (sitting systolic and diastolic blood pressure, pulse rate, and body temperature) are to be measured prior to blood sampling and prior to dosing, where applicable.

ⁱ Pregnancy testing will consist of serum testing at the Screening Visit. Urine pregnancy tests will be performed at all other visits where specified.

^j See Exclusion Criterion 9.

Table 1–1: Schedule of activities

Protocol activity	Visit Name or Weeks After 1 st Dose ^a	Screening ^b	Initial Treatment Period (weeks after first dose)									Maintenance Treatment Period (weeks after first dose)															PEOT	SFU ^c
			Bsl 1 st dose 0	1	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44		

^k The HIV test result will not be recorded in the eCRF.

^l A chest x-ray must be performed at the Screening Visit, or must occur within 2 months prior to Screening and results must be available at Baseline. A CAT scan of chest at Screening or within 2 months prior to Screening is acceptable, if available. The official report of the imaging must specifically indicate the absence of findings suggestive of current active TB or prior inactive TB.

^m Quantiferon Gold Plus will be used for this analysis. Details on IGRA assessment are provided in Section 8.2.6.3.4.

ⁿ Lesion count will be performed at the specified study visits, and must address all relevant anatomical regions in each study participant; Hurley Stage is included as an assessment performed during the lesion count. The data collected from the lesion count will be used for the derivation of the HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀, HS Physician's Global Assessment, and International Hidradenitis Suppurativa Severity score system (IHS4).

^o At certain sites, where feasible, representative photographs of the changes in skin will be captured. Photographs will be anonymized.

^p All blood samples taken prior to dosing.

^q All genomic, proteomic, and metabolomic samples and the genetic/epigenetic samples will be stored at -80°C at the central biorepository for up to 20 years.

^r The HSSDD (pain, smell or odor, drainage or oozing from HS lesions, and itch) will be completed daily from Screening through Week 16.

^s Randomization occurs at the Baseline Visit for all study participants.

^t IMP administration is based on randomization at the Baseline Visit for all study participants. Double-dummy IMP administration will occur every 2 weeks in all study participants to maintain the study blind. The dosing window is ± 3 days relative to the scheduled dosing visit.

2 INTRODUCTION

2.1 Study rationale

UCB is investigating bimekizumab (a humanized, full-length immunoglobulin [Ig] G1 monoclonal antibody [mAb] with high affinity for human interleukin [IL] 17 [IL-17A and IL 17F]) for the treatment of hidradenitis suppurativa (HS). Antibodies targeting IL-17A are effective in treating patients with moderate to severe psoriasis (PSO) and other immuno-inflammatory conditions. Data from clinical studies with bimekizumab suggest that inhibition of both IL-17A and IL-17F could be beneficial to patients with such conditions, including HS. Based on results from a Phase 2a study in study participants with HS (HS0001; European Union Drug Regulating Authorities Clinical Trials [EudraCT] Number 2017-000892-10, NCT03248531) and a Phase 2b study in study participants with PSO (PS0010) and its 48 week extension study (PS0011), bimekizumab doses of 320mg every 2 weeks (Q2W) in HS and 320mg every 4 weeks (Q4W) in PSO, appeared to have an acceptable safety profile, and achieved clinically meaningful efficacy in both primary and key secondary endpoints in their respective studies. Based on these data and considering the high unmet need for safe and effective therapies for HS, confirmatory studies are being conducted with bimekizumab for the treatment of moderate to severe HS.

2.2 Background

Hidradenitis suppurativa or acne inversa is a chronic, inflammatory, recurrent, debilitating skin disease that usually presents after puberty with painful, deep-seated, inflamed lesions in the apocrine gland-bearing areas of the body, most commonly the axillaries, inguinal, and anogenital regions (Dessau definition, First International Conference on HS, 30 Mar to 01 Apr 2006, Dessau, Germany). The nodules are often inflamed, can progress to abscess formation, and may rupture to form fistulas and subsequent scarring. Thus, many patients with HS develop permanent sequelae of past inflammation that are only remediable through surgical excision of the involved skin areas. Hidradenitis suppurativa is also associated with several complications (eg, the development of anal, urethral, and rectal strictures and fistulas), and the excessive scarring and fibrosis produced by HS lesions can lead to contractures and limitations in limb mobility (Alikhan, 2009; Danby, 2010).

Hidradenitis suppurativa is estimated to affect about 1% of the adult European population, with a female to male ratio of approximately 3:1 (Revuz, 2008; Naldi, 2006). The prevalence of diagnosed HS in the US may be lower (<0.1%), although further research is needed to determine the prevalence of undiagnosed HS in the US (McMillan, 2014). Patients diagnosed with HS often experience a significant reduction in quality of life (QOL), equivalent to severe PSO (Sartorius, 2009), due to the location of, and discharge from, the lesions that leads to an often persistent morbidity due to pain and sequelae from uncontrolled inflammation (von der Werth, 2001; Wolkenstein, 2007). The reduction in QOL and persistent morbidity result in functional impairment in patients with HS similar or greater to that of heart disease, diabetes, or asthma, when measured by the European Quality-of-Life 5 dimensions 3-level questionnaire (EQ-5D-3L) scale (Riis, 2016).

Bimekizumab is a humanized full-length mAb of IgG1 subclass being developed for the treatment of patients with inflammatory diseases such as PSO, psoriatic arthritis, axial spondyloarthritis, and HS. Bimekizumab has high affinity for human IL-17A and human IL-17F,

and selectively and potently inhibits the activity of both isoforms in vitro. The key pro-inflammatory cytokine IL-17A has been demonstrated to, and IL-17F is believed to, play important roles in autoimmune and inflammatory diseases. Published data and immunohistochemistry studies performed by UCB have shown that expression of both IL-17A and IL-17F is present in HS lesions, and there are published reports highlighting the potential for IL-17A and IL-17F to drive HS disease pathology (UCB Research Report 40001864; [Cho](#), 2012; [Schlapbach](#), 2011). This supports the hypothesis that the IL-17 cytokine family is a potential therapeutic target in HS. Bimekizumab is hypothesized to demonstrate a treatment response in HS because it selectively and potently inhibits the activity of both IL-17A and IL-17F isoforms in vitro.

2.3 Benefit/risk assessment

Overall, the Phase 2a POC study in HS (HS0001) demonstrated that the safety profile (during 10 weeks of dosing for a 12-week treatment period) for bimekizumab 320mg sc Q2W appears consistent with that of bimekizumab in other indications (PSO, psoriatic arthritis, axial spondyloarthritis) for which bimekizumab is being developed.

Treatment-emergent adverse events (TEAEs) experienced by study participants with HS receiving repeated doses of bimekizumab occurred at incidences similar to those with placebo and adalimumab (range: 61.9% to 71.4%). In both the bimekizumab and adalimumab groups, the most frequently reported TEAEs were in the system organ classes of Infections and infestations (43.5% and 42.9%, respectively), Skin and subcutaneous tissue disorders (28.3% and 42.9%, respectively), and General disorders and administration site conditions (21.7% and 23.8%, respectively). The most frequently reported TEAEs in the placebo group were in the SOC of Nervous system disorders (28.6%) and Infections and infestations and Skin and subcutaneous tissue disorders (19.0% each). The most frequently reported TEAEs (by preferred term) were hidradenitis (17.4%) and fatigue (8.7%) in the bimekizumab group, hidradenitis (33.3%) and influenza (14.3%) in the adalimumab group, and headache and hidradenitis (14.3% each) and arthralgia (4.8%) in the placebo group. Events of hidradenitis were not unexpected in a population of study participants with moderate-to-severe HS.

With regard to TEAEs of special interest and other safety topics of interest, there were no major adverse cardiovascular events, serious fungal/opportunistic infections (including tuberculosis [TB]), malignancies (including lymphoma), neuropsychiatric events, cases of inflammatory bowel disease (IBD), evidence of hepatotoxicity (per Hy's Law criteria), or hypersensitivity/anaphylactic reactions reported with bimekizumab treatment. No new safety signals were identified in HS0001 compared with those observed with bimekizumab across other development programs to date.

No clinically relevant patterns of changes were observed in any treatment group in hematology, clinical chemistry, vital signs, or electrocardiogram (ECG) findings. Few markedly abnormal post-Baseline liver function test values were reported during the study.

With respect to benefit, the totality of the data from HS0001 demonstrated that bimekizumab results in clinically meaningful and consistent improvements versus placebo across all HS outcome measures assessed. Improvements began early after initiation of treatment, and persisted through the last assessment (Week 12). The efficacy data for the primary endpoint, HiSCR₅₀ (ie, a 50% reduction in the total abscess and inflammatory nodule [AN] count with no

increase from Baseline in abscess or draining tunnel count) at Week 12, were comparable with that of adalimumab, and other efficacy measures suggested improved therapeutic outcomes (larger proportions of study participants achieving HiSCR₇₅ [a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count], HiSCR₉₀ [a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count], and Dermatology Life Quality Index [DLQI] total scores of 0 and 1).

Overall, results of HS0001 show that bimekizumab appears to have an acceptable safety profile considering the anticipated benefit at a dose of 320mg Q2W for the treatment duration evaluated to date. No new safety concerns were raised in study participants with moderate to severe HS, and the benefit/risk remains positive and supports continued investigation of bimekizumab at 320mg Q2W.

More detailed information about the known and expected benefits and risks of bimekizumab may be found in the Investigator's Brochure.

3 OBJECTIVES AND ENDPOINTS

Objectives	Endpoints
Primary	
Evaluate the efficacy of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₅₀ at Week 16
Secondary	
Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS	<ul style="list-style-type: none"> HiSCR₇₅ at Week 16 Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16 Absolute change from Baseline in DLQI Total Score at Week 16 Absolute change from Baseline in Skin Pain score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HSSDD Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) at Week 16
Evaluate the safety of bimekizumab in study participants with moderate to severe HS	<ul style="list-style-type: none"> Treatment-emergent AEs Serious TEAEs TEAEs leading to withdrawal from study
Other	
Evaluate the efficacy of bimekizumab on HiSCR, other HS Scores, and other clinical measures of disease activity at various timepoints in study participants with moderate to severe HS	<ul style="list-style-type: none"> Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ Absolute change from Baseline in International Hidradenitis Suppurativa Severity Score System Change from Baseline in the HS-Physician’s Global Assessment 6-point scale Absolute and percentage change from Baseline in hs-CRP Initiation of systemic antibiotic rescue therapy HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ at both Weeks 16 and 48 Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders

Objectives	Endpoints
	<ul style="list-style-type: none"> Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₅₀ during the Maintenance Treatment Period Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period
Evaluate the efficacy of bimekizumab on abscesses, nodules, and draining tunnels at various timepoints in study participants with moderate to severe HS	<ul style="list-style-type: none"> Change and percentage change from Baseline in lesion counts (abscess count, inflammatory nodule count, AN count, and draining tunnel count) AN count of 0, 1, or 2 AN₂₅, AN₅₀, AN₇₅, AN₉₀, AN₁₀₀ Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48 Time to flare from Weeks 0 to 16 Time to flare from Week 16 to 48
Evaluate the efficacy of bimekizumab on patient-reported outcome measures at various timepoints in study participants with moderate to severe HS	<ul style="list-style-type: none"> Absolute and percentage change (worst and average pain) from Baseline in HS Skin Pain score (11-point numeric rating scale) Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) Pain response (defined as a decrease from Baseline in HSSQ weekly worst skin pain score at or beyond the threshold for clinically meaningful change) Pain response (at least a 30% reduction and at least a 1-unit reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline Pain response (at least a 30% reduction and at least a 1-unit reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline Pain response (at least a 30% reduction and at least a 1-unit reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline Absolute change from Baseline in DLQI Total Score DLQI Total Score of 0 or 1

Objectives	Endpoints
	<ul style="list-style-type: none"> • Minimum clinically important difference (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4) • Absolute change from Baseline in HiSQOL domain scores (symptoms, psychosocial, activities, adaptations) • Absolute change from Baseline in Patient Global Impression of HS Severity • Absolute change from Baseline in Patient Global Impression of Severity of HS Skin Pain • Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor. • Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor • Responses to the EQ-5D-3L, absolute and changes from Baseline in EQ-5D-3L visual analog scale scores • Absolute change from Baseline in Work Productivity and Activity Impairment Questionnaire-Specific Health Problem v2.0 adapted to HS scores • Domain Scores (effectiveness, convenience and global satisfaction) on the Treatment Satisfaction Questionnaire – Medication-9
<p>Evaluate the effect of bimekizumab on other safety measures at various timepoints in study participants with moderate to severe HS</p>	<ul style="list-style-type: none"> • Adverse events of special interest (Hy's Law) • Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behavior, , major adverse cardiovascular events, hepatic events and potential drug-induced liver injury (PDILI), malignancies, and inflammatory bowel disease. • Absolute change from Baseline in the PHQ-9 score • Absolute change from Baseline in vital signs • Absolute change from Baseline in clinical laboratory values (chemistry and hematology) • ECG results
<p>Evaluate the pharmacokinetics of bimekizumab in study participants with moderate to severe HS</p>	<ul style="list-style-type: none"> • Plasma bimekizumab concentrations over the study duration

Objectives	Endpoints
Evaluate the immunogenicity of bimekizumab (antidrug antibodies) in study participants with moderate to severe HS	<ul style="list-style-type: none"> Bimekizumab antidrug antibodies Bimekizumab neutralizing antibodies
Exploratory	
Evaluate biomarkers in study participants with moderate to severe HS.	<ul style="list-style-type: none"> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <p>The candidate exploratory biomarkers are the blood or blood derivative (eg, plasma) concentrations of</p> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div> <div style="background-color: black; height: 1.2em; width: 100%;"></div>

AE=adverse event; AN=abscess and inflammatory nodule; AN₂₅=a 25% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₅₀=a 50% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₇₅=a 75% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₉₀=a 90% reduction in the total abscess and inflammatory nodule count relative to Baseline; AN₁₀₀=a 100% reduction in the total abscess and inflammatory nodule count relative to Baseline; DLQI=Dermatology Life Quality Index; ECG=electrocardiogram; eC-SSRS=electronic Columbia Suicide Severity Rating Scale; EQ-5D-3L=Euro-Quality of Life 5-Dimensions, 3 levels; HiSCR=Hidradenitis Suppurativa Clinical Response; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₉₀=a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₁₀₀=a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSQOL=Hidradenitis Suppurativa Quality of Life; HS=hidradenitis suppurativa; hs-CRP=high-sensitivity C-reactive protein; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; HSSQ=Hidradenitis Suppurativa Symptom Questionnaire; PDILI=potential drug-induced liver injury; PHQ-9=Patient Health Questionnaire Depression Module; QOL=quality of life; TB=tuberculosis; TEAE=treatment-emergent adverse event

4 STUDY DESIGN

4.1 Overall design

HS0004 is a Phase 3, randomized, double-blind, placebo-controlled, multicenter, pivotal study evaluating the efficacy and safety of bimekizumab in study participants with moderate to severe HS. Study participants meeting the inclusion criteria who do not meet any exclusion criteria will complete a Screening Period of 14 days to up to 5 weeks; a double-blind, 48-week Treatment Period comprising a 16-week Initial Treatment Period and 32-week Maintenance Treatment Period; and a 20-week Safety Follow-up (SFU) Period following the final injection of investigational medicinal product (IMP) if study participants do not enter a subsequent extension study (HS0005) or withdraw prematurely from treatment.

Study participants will be randomized in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to 1 of 3 dose regimens of bimekizumab or placebo as shown in the schematic (Figure 1-1). All doses of IMP will be administered by sc injection. The primary efficacy variable at Week 16 is HiSCR₅₀. Study visits will occur at Screening; Baseline (Week 0); Weeks 1, 2, 4, 6, 8, 10, 12, 14, and 16; and every 2 weeks from Week 16 through Week 48 for assessments of efficacy, safety, and other measures of QOL/health status/work productivity. An SFU visit will be conducted 20 weeks after the last dose of IMP for participants who do not enter the extension study, or who are prematurely withdrawn from the study.

4.1.1 Screening Period (Weeks -5 to 0)

The Screening Period will last a minimum of 14 days up to a maximum of 5 weeks prior to randomization.

4.1.2 Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)

Eligible study participants will be randomized in a 2:2:2:1 ratio as noted in the Study Schema to:

- Bimekizumab 320mg Q2W from Weeks 0 to 48
- Bimekizumab 320mg Q4W from Weeks 0 to 48
- Bimekizumab 320mg Q2W to Week 16, continuing on 320mg Q4W from Weeks 16 to 48
- Placebo to Week 16, continuing on bimekizumab 320mg Q2W from Weeks 16 to 48

4.1.3 Safety Follow-up Visit

All study participants not continuing in the extension study, including those withdrawn from IMP as well as those completing all study visits, will have an SFU Visit 20 weeks after their final dose of IMP.

4.1.4 Visit Windows

Visit windows are ± 3 days (based on the date of the first dose). The minimum number of days between 2 consecutive injection visits is 8 days (eg, Visit 2 +3 days occurs on Day 17; Visit 4 -3 days occurs on Day 25). However, the minimum of 8 days between injections may be used only 1 time in the Initial Treatment Period and 1 time in the Maintenance Treatment Period, if needed. The study participant should be dosed according to the administration schedule thereafter. The 20-week SFU Visit window is ± 7 days (based on the date of the final dose).

4.2 Scientific rationale for study design

A randomized, double-blind, placebo-controlled study design has been selected to demonstrate efficacy and safety of bimekizumab for regulatory approval. The study population will include adults with moderate to severe HS. The inclusion and exclusion criteria were designed to ensure the safety of study participants, and to enroll a broad HS study participant population representative of clinical practice in terms of disease severity and morbidity (physical disability and discomfort) that warrants therapy with a systemic agent. Considering that study participants with moderate to severe HS are treated with different antibiotics, systemic tetracyclines have been selected as the most appropriate class of antibiotics for the study based on the current therapeutic guidelines for HS.

The primary efficacy outcome measure (HiSCR₅₀) is a validated clinical outcome measure for evaluating efficacy in study participants with moderate to severe HS.

In addition, a core domain set for HS study outcome established for HS calls for the concurrent measurement of 5 core outcome domains agreed by both patients and health care providers: pain, physical signs, HS specific quality of life, global assessment and progression of course. A sixth domain, symptoms, has also been recommended by the Steering Group because it received strong support from the patient stakeholder group (Thorlacius, 2018).

The Screening Period is included to ensure eligibility criteria are met, including collection of laboratory data, verification that the doses of concomitant and allowable medications are stable, and to enable washout of any medications not permitted for use during the study.

The randomization allocation and sample sizes have been selected to (1) maximize exposures to bimekizumab test doses/regimens, (2) ensure adequate power to demonstrate superiority of bimekizumab to placebo for the primary endpoint (HiSCR₅₀) at Week 16, and (3) have sufficient sample size to detect statistically significant differences between treatments as specified in the secondary endpoints at Week 16.

An initial treatment period of 16 weeks will be used to demonstrate the efficacy of bimekizumab over placebo. The 32-week Maintenance Treatment Period will collect information on safety and efficacy beyond initial treatment.

The Maintenance Treatment Period is designed to assess the durability of response of the study endpoints and to provide sufficient longer-term (up to 48 weeks, including the Initial Treatment Period) safety and exposure to bimekizumab for regulatory filings. Continuous (up to 48 weeks) exposure to both the 320mg bimekizumab Q2W and 320mg bimekizumab Q4W dose regimens allow for the following assessments:

- Durability of response and longer-term safety
- Optimal dosing interval for maintenance treatment

This period will also allow study participants who received placebo in the Initial Treatment Period to begin receiving bimekizumab at Week 16 in the randomized, controlled Maintenance Treatment Period of the study.

4.3 Justification for dose

The pathophysiology of HS is an active area of research, with investigations targeting identification of the cytokines and immune pathways in HS. Recent reviews on the subject indicate the potential diversity of these inflammatory pathways and mediators, and the impact on treatment response to various pharmacologic interventions. As concluded by Frew, Hawkes, and Krueger, no current schema accurately predicts treatment efficacy to date (Prens, 2015; Frew, 2018). Furthermore, the inflammatory burden in HS seems greater than other autoinflammatory conditions affecting the skin (Van der Zee, 2011; Riis, 2015, Martorell, 2015). These data, and the current pharmacologic treatment of HS suggest that a more intensive dosing regimen (ie, dose level and/or frequency of administration) may be needed for the treatment of HS.

Consistent with the above literature, a 320mg Q2W dose regimen of bimekizumab is being evaluated in the Phase 3 program for HS. This dose regimen is higher than those used in other

indications currently in Phase 3 development for bimekizumab; however, HS0001 results revealed that bimekizumab 320mg Q2W demonstrated consistent, clinically meaningful efficacy in the treatment of HS when compared to placebo, with safety results consistent with those of studies of bimekizumab in other indications in development. In addition, pharmacokinetic (PK) data from HS0001 demonstrated that study participants with HS have a lower exposure to bimekizumab than study participants with PSO given the same dose and regimen, thus necessitating higher bimekizumab doses for HS (maximum monthly bimekizumab dose of 640mg).

In addition to the 320mg Q2W dose used in HS0001, the highest dose being used in other bimekizumab indications (PSO dose regimen of 320mg Q4W) is being used in this study. The assessment of both the Q2W and Q4W dosing frequencies in this study will help determine the optimal monthly bimekizumab dose required to sustain efficacy with long-term (maintenance) treatment.

4.4 End of study definition

A study participant will be considered to have completed the study if he or she completed the Week 48 visit.

The end of the study is defined as the date of the last scheduled procedure shown in the Schedule of Activities ([Table 1–1](#)) for the last study participant in the study globally, including the SFU, as applicable.

5 STUDY POPULATION

Prospective approval of protocol deviations to recruitment and enrollment criteria, also known as protocol waivers or exemptions, is not permitted.

5.1 Inclusion criteria

Participants are eligible to be included in the study only if all of the following criteria apply:

Age

1. Participant must be at least 18 years of age, at the time of signing the informed consent. If a study participant is under the local age of consent and is at least 18 years of age, written informed consent will be obtained from both the study participant and the legal representative.

Type of participant and disease characteristics

- 2a. Study participants must have a diagnosis of HS based on clinical history and physical examination for at least 6 months prior to the Baseline visit; diagnosis must be verifiable through medical notes and documentation.
3. Study participant must have HS lesions present in at least 2 distinct anatomic areas (eg, left and right axilla), 1 of which must be at least Hurley Stage II or Hurley Stage III at both the Screening and Baseline visits.
4. Study participant must have moderate to severe HS defined as a total of ≥ 5 inflammatory lesions (ie, number of abscesses plus number of inflammatory nodules) at both the Screening and Baseline visits.

5a. Study participant must have had a history of inadequate response to a course of a systemic antibiotics for treatment of HS at the Screening Visit as assessed by the Investigator through study participant interview and review of medical history; inadequate response must be verifiable through medical notes and documentation. Study participants who meet any of the following are NOT automatically excluded from the study:

- Demonstrated intolerance to (or during therapy became intolerant to) systemic antibiotics
- Had a contraindication to systemic antibiotics
- Responded to course(s) of systemic antibiotic(s) and subsequently exhibited recurrence after discontinuation of the antibiotic

Sex

6. Males and females may be study participants.

- A female study participant is eligible to participate if she is not pregnant (see Appendix 4 [Section 10.4]), not breastfeeding, and at least one of the following conditions applies:
 - Not a woman of childbearing potential (WOCBP) as defined in Appendix 4
 - OR
 - A WOCBP who agrees to follow the contraceptive guidance in Appendix 4 during the treatment period and for at least 20 weeks after the last dose of IMP.

Informed consent

7. Study participant was capable of giving signed informed consent as described in Appendix 1 which includes compliance with the requirements and restrictions listed in the informed consent form (ICF) and in this protocol.

5.2 Exclusion criteria

Study participants are excluded from the study if any of the following criteria apply:

Medical conditions

1. Study participant has any medical or psychiatric condition that, in the opinion of the Investigator, could jeopardize or would compromise the study participant's ability to participate in this study as determined by the Investigator based on protocol-required assessments.

HS, Skin-Specific, and Other Inflammatory Disease

2. Study participant has a draining tunnel count of >20 at the Baseline Visit.
3. Study participant has any other active skin disease or condition (eg, bacterial cellulitis, candida intertrigo, extensive condyloma) that may, in the opinion of the Investigator, interfere with the assessment of HS.
4. Study participant has a diagnosis of sarcoidosis, systemic lupus erythematosus, or active IBD. Note: Study participants with a diagnosis of Crohn's disease or ulcerative colitis are allowed if they have no active symptomatic disease at Screening or Baseline.

5. Study participant has a primary immunosuppressive condition, including taking immunosuppressive therapy following an organ transplant, or has had a splenectomy.

Other Medical Conditions

6. Female study participant who is breastfeeding, pregnant, or plans to become pregnant during the study or within 20 weeks following the final dose of IMP.

7. Study participant has an active infection or history of infection(s) as follows:

- Any infection requiring systemic treatment within 14 days prior to Baseline
- A serious infection, defined as requiring hospitalization or intravenous anti-infective(s) within 2 months prior to the Baseline Visit
- A history of opportunistic, recurrent, or chronic infections that, in the opinion of the Investigator, might cause this study to be detrimental to the study participant. Opportunistic infections are infections caused by uncommon pathogens (eg, *Pneumocystis jirovecii*, cryptococcosis), or unusually severe infections caused by common pathogens (eg, cytomegalovirus, herpes zoster)

8. Study participant has any of the following:

- Known active TB disease
- History of active TB involving any organ system unless adequately treated according to World Health Organization/Centers for Disease Control and Prevention therapeutic guidance and proven to be fully recovered upon consult with a TB specialist
- Latent TB infection (LTBI). Participants with LTBI diagnosed during Screening must have completed a course of prophylaxis prior to IMP dosing. Participants can be rescreened after completion of a full course of prophylaxis plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline to avoid any interference with the study efficacy measurements (eg, concomitant antibiotics). Prophylaxis should be in accordance with applicable clinical guidelines and TB specialist judgment based on the origin of infection.
- High risk of exposure to TB infection
- Current pulmonary nontuberculous mycobacterial (NTM) infection or history of pulmonary NTM infection unless proven to be fully recovered

Note: For further information relating to definitions of known active TB, past history of TB, LTBI, high risk of acquiring TB infection and NTM infection refer to Section 8.2.6.

9. Study participant has an acute or chronic hepatitis B virus, hepatitis C virus (HCV), or human immunodeficiency virus (HIV) infection. Study participants who have evidence of, or tested positive for, hepatitis B or hepatitis C will be excluded. A positive test for hepatitis B virus is defined as: 1) positive for hepatitis B surface antigen, or 2) positive for anti-hepatitis B core antibody. A positive test for HCV is defined as: 1) positive for hepatitis C antibody, and 2) positive via a confirmatory test for HCV (for example, HCV polymerase chain reaction).

10. Study participants with concurrent malignancy are excluded. Study participants with a history of malignancy within the past 5 years prior to the Screening Visit are excluded, EXCEPT if the malignancy was a cutaneous squamous or basal cell carcinoma, or in situ cervical cancer that has been treated and is considered cured.
11. Study participant has a history of a lymphoproliferative disorder including lymphoma or current signs and symptoms suggestive of lymphoproliferative disease.
12. Study participant has had major surgery within the 3 months prior to the Baseline Visit, or has planned major surgery after entering the study.
13. Study participant has any systemic disease (ie, cardiovascular, neurological, renal, liver, metabolic, gastrointestinal, hematological, immunological, etc.) considered by the Investigator to be uncontrolled, unstable, or likely to progress to a clinically significant degree during the course of the study.
14. Study participant has had a myocardial infarction or stroke within the 6 months prior to the Screening Visit.
15. Study participant has a history of chronic alcohol or drug abuse within 6 months prior to Screening as evaluated by the Investigator based on medical history, interview, and/or results of the Screening urine drug screen.
- 16a. Study participant has the presence of active suicidal ideation, or positive suicide behavior using the “Screening” version of the electronic Columbia Suicide Severity Rating Scale (eC-SSRS) with either of the following criteria:
 - Study participant has a history of a suicide attempt within the 5 years prior to the Screening Visit. Study participants with a history of a suicide attempt more than 5 years ago should be evaluated by a mental healthcare practitioner before enrolling into the study.
 - Suicidal ideation in the past month prior to the Screening Visit as indicated by a positive response (“Yes”) to either Question 4 or Question 5 of the “Screening” version of the eC-SSRS.
17. Study participant has presence of moderately severe major depression or severe major depression indicated by a score of ≥ 15 using the screening Patient Health Questionnaire Depression Module (PHQ-9). Medication used to treat depression should be stable for 8 weeks prior to Baseline.
18. Study participant has a known hypersensitivity to any components of bimekizumab or comparative drugs as stated in this protocol.

Prior/Concomitant therapy

- 19a. Study participant has had prior treatment with an IL-17 biologic response modifier or has participated in IL-17 biologic response modifier study unless an appropriate washout has been performed since the last dose of IMP (within 6 months prior to the Baseline Visit or 5 half-lives [whichever is greater]).
20. Study participant received prescription topical therapies for the treatment of HS within 14 days prior to the Baseline Visit.

21. Study participant is currently receiving systemic nonbiologic or biologic therapies for HS with potential therapeutic impact for HS. Note: If study participant received systemic nonbiologic or biologic therapies for HS and stopped these treatments, washout periods should be applied as shown in Table 6–3. Note: this does not apply to study participants who may be eligible for randomization into the antibiotic strata.
22. If study participant is using concomitant, non-opioid analgesics for HS-related or non-HS-related pain as permitted by protocol, they should be on a stable (scheduled) dose for at least 14 days prior to the Baseline Visit and anticipate continuing that dose through Week 16 unless a decrease in dose is warranted based on symptoms. Opioid analgesics are excluded. Note: As needed (PRN) use is not considered a stable dose, but (for example) taking a nonsteroidal anti-inflammatory drug (NSAID) 3 times per week, every week is considered a stable dose.
23. Study participant has received any live (including attenuated) vaccination within the 8 weeks prior to the Baseline Visit (eg, inactivated influenza and pneumococcal vaccines are allowed, but nasal influenza vaccination is not permitted). Live vaccines are not allowed during the study, including the SFU Period (20 weeks after the last dose of IMP).
24. Study participant has received Bacillus Calmette-Guerin vaccination within 1 year prior to IMP administration.

Prior/Concurrent clinical study experience

25. Study participant has previously participated in this study or study participant has previously been assigned to treatment in a study of the medication under investigation in this study, and received at least 1 dose of IMP (including placebo).
26. Study participant is currently participating in another study of a systemic medication under investigation, including SFU. Study participant must be washed out of the medication as indicated in Table 6–3.
27. Study participant is currently participating in another study of a topical medication under investigation, including SFU. Study participant must be washed out of the medication for 4 weeks prior to the Baseline Visit.
28. Study participant is currently, or was within the 4 weeks prior to the Baseline Visit, participating in another study of a medical device under investigation.

Diagnostic assessments

29. Study participant has laboratory abnormalities at Screening, including any of the following:
 - $\geq 3 \times$ the upper limit of normal (ULN) of any of the following: alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP)
 - Bilirubin $> 1.5 \times \text{ULN}$ (isolated bilirubin $> 1.5 \times \text{ULN}$ is acceptable if bilirubin is fractionated and direct bilirubin $< 35\%$)
 - White blood cell count $< 3.00 \times 10^3/\mu\text{L}$
 - Absolute neutrophil count $< 1.5 \times 10^3/\mu\text{L}$

- Lymphocyte count <500 cells/ μ L
- Hemoglobin <8.5g/dL

Note: Individual screening tests for which the results are in error, borderline, or indeterminate for inclusion in the study can be repeated once for confirmation during the Screening Period. Upon retesting, study participants whose results remain outside this threshold should not be randomized.

30. Study participant has any other laboratory abnormality, which, in the opinion of the Investigator, will prevent the study participant from completing the study or will interfere with the interpretation of the study results.

Other exclusions

31. Study participant is a UCB employee or is an employee of third-party organizations involved in the study.
32. Study participant and/or his or her immediate family member is an employee, volunteer, or other worker at the investigative site either affiliated or not affiliated with this study. Immediate family is defined as a spouse, parent, child, or sibling whether biological or legally adopted.

5.3 Lifestyle restrictions

Not applicable to this study.

5.4 Screen failures

Screen failures are defined as participants who consent to participate in the clinical study but are not subsequently randomized to study treatment. A minimal set of screen failure information is required to ensure transparent reporting of screen failure participants to meet the Consolidated Standards of Reporting Trials (CONSORT) publishing requirements and to respond to queries from regulatory authorities. Minimal information includes demography, screen failure details, eligibility criteria, and any serious adverse events (SAE). Individuals who do not meet the criteria for participation in this study (screen failure) may be rescreened, following discussion with the Medical Monitor or Sponsor's study physician.

Participants who are rescreened should be assigned a new participant number for rescreening.

A study participant may be rescreened 1 time for reasons including, but not limited to, the following:

- Individual laboratory screening tests for which the results are exclusionary can be retested (eg, tests that result in ALT, AST, or ALP up to 25% above the exclusion limit may be repeated once for confirmation. Test can also be repeated during rescreening.) Of note, repetition of laboratory screening tests within the Screening Period is permitted for technical reasons (eg, frozen sample, expired laboratory kit) without contacting the Medical Monitor.
- Eligibility assessments that could not be completed as planned (eg, for technical reasons) within the defined Screening Period of 5 weeks without approval by Medical Monitor.
- Abnormal ECG results.

- Did not meet the required washout period for concomitant medications.
- Study participant needs to complete a full course of antibiotic therapy for LTBI plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline as described in Exclusion Criterion 8 (Section 5.2).
- If the study participant requires an incision and drainage procedure for a HS lesion(s) during the Screening Period, the study participant should be screen failed. The participant can be rescreened when the lesion is considered healed. The study participant must have completed antibiotics/analgesic treatment if required for the procedure before rescreening as described in Table 6–3.

Study participants who fail to meet the eligibility criteria for PHQ-9, eC-SSRS, or the TB questionnaire are not allowed to be rescreened.

The Medical Monitor must be contacted for confirmation of rescreening/retesting in all other cases.

6 STUDY TREATMENTS

Study treatment is defined as any investigational treatment, marketed product, placebo, or medical device intended to be administered to a study participant according to the study protocol.

6.1 Treatments administered

Eligible study participants will be randomized in a 2:2:2:1 ratio as noted in the Study Schema (Figure 1-1).

A summary of the treatments administered is provided in Table 6–1.

Table 6–1: Study medications administered

ARM Name	Bimekizumab	Placebo
Intervention name	Bimekizumab	Placebo
Type	Biologic	Drug
Dose formulation	Solution for injection (pre-filled 1-mL syringe)	Solution for injection (pre-filled 1-mL syringe)
Unit dose strengths	160mg/mL	0.9% sodium chloride aqueous solution (physiological saline, preservative free) of US Pharmacopoeia/European Pharmacopoeia quality appropriate for injection; same volume to maintain blinding to the study participant
Dosage levels	320mg	Not applicable
Route of administration	Subcutaneous injection	Subcutaneous injection
Use	Experimental	Placebo comparator
Investigational Medicinal Product and Non-Investigational Medicinal Product	Investigational Medicinal Product	Investigational Medicinal Product
Sourcing	Provided centrally by the Sponsor.	Provided centrally by the Sponsor.
Packaging and labeling	Study Intervention will be provided in a kit. Each kit will be labeled as required per country requirement	Study Intervention will be provided in a kit. Each kit will be labeled as required per country requirement
Current/Former names or aliases	Bimekizumab	Not applicable

Because of differences in the dosing schedules and in order to maintain blinding, all study participants will receive 2 injections sc every 2 weeks from Week 0 to Week 46 as depicted in [Table 6–2](#).

Table 6–2: Dosing scheme

Week Dose Assignment	Initial Treatment Period (weeks after first dose)										Maintenance Treatment Period (weeks after first dose)															
	Baseline 0	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46		
Bimekizumab 320mg Q2W/Q2W	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●	●●		
Bimekizumab 320mg Q4W/ Q4W	●●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○		
Bimekizumab 320mg Q2W/ 320mg Q4W	●●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○	● ●	○ ○		
Placebo/bimekizumab 320mg Q2W	○○	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	○ ○	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●	● ●		

Q2W=every 2 weeks; Q4W=every 4 weeks

Notes: A bimekizumab 160mg injection is depicted by a black circle (●). A placebo injection is depicted by a white circle (○).

6.2 Preparation, handling, storage, and accountability requirements

The Investigator or designee must confirm appropriate temperature conditions have been maintained during transit for all study treatment received and any discrepancies are reported and resolved before use of the study treatment.

Only study participants enrolled in the study may receive study treatment and only authorized site staff may supply or administer study treatment. All study treatments must be stored in a secure, environmentally controlled, and monitored (manual or automated) area in accordance with the labeled storage conditions with access limited to the Investigator and authorized site staff.

The Investigator, institution, or the head of the medical institution (where applicable) is responsible for study treatment accountability, reconciliation, and record maintenance (ie, receipt, reconciliation, and final disposition records).

In case an out-of-range temperature is noted, it must be immediately reported as per instructions contained in the IMP Handling Manual.

Further guidance and information for the final disposition of unused study treatment are provided in the IMP Handling Manual.

6.2.1 Drug accountability

A Drug Accountability form will be used to record IMP dispensing and return information on a by-participant basis and will serve as source documentation during the course of the study. Details of any IMP lost, damaged (due to breakage or wastage), not used, partially used, disposed of at the study site, or returned to the Sponsor or designee must also be recorded on the appropriate forms. All supplies and pharmacy documentation must be made available throughout the study for UCB (or designee) to review.

The Investigator (or designee) is responsible for retaining all used, unused, and partially used containers of IMP until returned or destroyed.

The Investigator may assign some of the Investigator's duties for drug accountability at the study site to an appropriate pharmacist/designee.

The Investigator must ensure that the IMP is used only in accordance with the protocol.

Periodically, and/or after completion of the clinical phase of the study, all used (including empty containers)/partially used, unused, damaged, and/or expired IMP must be reconciled and either destroyed at the site according to local laws, regulations, and UCB standard operating procedures or returned to UCB (or designee). Investigational medicinal product intended for the study cannot be used for any other purpose than that described in this protocol.

6.3 Measures to minimize bias: randomization and blinding

6.3.1 Randomization and numbering of participants

An interactive response technology (IRT) system will be used for assigning eligible study participants to a treatment regimen (as applicable) based on a predetermined production randomization and/or packaging schedule provided by UCB (or designee). The randomization schedule will be produced by the IRT vendor. The IRT will generate individual assignments for kits of IMP, as appropriate, according to the visit schedule.

To enroll a study participant at Screening, the Investigator or designee will contact the IRT and provide brief details about the study participant to be enrolled. Each study participant will receive a unique number assigned at Screening that serves as the study participant identifier throughout the study. The study participant number will be required in all communication between the Investigator or designee and the IRT regarding a particular study participant. Study participant numbers and kit numbers will be tracked via the IRT.

To randomize a study participant, the Investigator or designee will contact the IRT and provide brief details about the study participant to be randomized. The IRT will automatically inform the Investigator or designee of the study participant's randomization number. The IRT will allocate kit numbers to the study participant based on the study participant number during the course of the study. The randomization number must be incorporated into the case report form (CRF).

6.3.2 Procedures for maintaining and breaking the treatment blind

6.3.2.1 Maintenance of study treatment blind

All study participant treatment details will be allocated and maintained by the IRT system.

The IRT provider will receive the randomization code at the start of the study.

Due to differences in presentation between bimekizumab and placebo treatments, special precautions will be taken to ensure study blinding; study sites will have blinded and unblinded personnel. Bimekizumab and placebo injections will be administered at the investigational sites by unblinded, dedicated study personnel according to the site-specific blinding plan. Unblinded study personnel will be responsible for recording the administration information on source documents, and administration of the IMP as sc injections. Study site pharmacists or other suitably qualified site personnel who are responsible for preparation and administration of IMP treatments will have access to treatment allocations for individual study participants via the IRT. The unblinded pharmacy monitors from the Contract Research Organization (CRO), and the UCB Clinical Trial Supply representative will also have access to the treatment allocations and to the drug accountability records, if applicable.

The following individuals may, as necessary, have access to the randomization code as indicated:

- Members of the Data Monitoring Committee (DMC) who participate in unblinded sessions will be given information about the IMP allocation for those study participants for whom data are provided.
- The unblinded, independent CRO staff supporting preparation of the data outputs for the DMC reviews.

The unblinded study site personnel will not be involved in the study in any way other than assuring the IMP is taken from the correct kit and prepared according to the IMP-handling manual, and administering the IMP to the study participants.

In addition, high-sensitivity C-reactive protein (hs-CRP) results will not be reported to any blinded study personnel as long as the study remains blinded.

6.3.2.2 Breaking the treatment blind in an emergency situation

In the event of an emergency, it will be possible to determine to which treatment arm and dose the study participant has been allocated by contacting the IRT. The Investigator is responsible for

breaking the treatment blind in case of emergency. All sites will be provided with details of how to contact the system for code breaking at the start of the study. The Medical Monitor and/or UCB study physician or equivalent should be consulted prior to unblinding, whenever possible.

The Clinical Project Manager will be informed immediately via the IRT when a code is broken, but will remain blinded to specific treatment information. Any unblinding of the IMP performed by the Investigator must be recorded in the source documents and on the Study Termination electronic CRF (eCRF) page.

Inadvertent unblinding will be listed as a major protocol deviation.

6.4 Treatment compliance

During the double-blind Initial Treatment and Maintenance Treatment Periods, IMP will be administered in the clinic and compliance will be recorded at the visit by study personnel in the eCRF. Drug accountability must be recorded on the Drug Accountability form.

6.5 Concomitant medications/treatments

6.5.1 Permitted concomitant treatments (medications and therapies)

The following concomitant medications are permitted during the study:

- Wound care: Concomitant use of wound care dressings on HS wounds is allowed; however, options are limited to alginates, hydrocolloids, and hydrogels and use of these will be recorded in the eCRF.
- Lesion care: Concomitant use of saline, water, and/or Vaseline (petroleum jelly) is allowed for care of skin lesions and use of these will be recorded in the eCRF.
- Analgesic therapy:
 - Study participants will be required to wash out of all analgesics for HS-related pain 14 days prior to Baseline. However, if a study participant is on a stable (scheduled) dose of a non-opioid analgesic for HS-related pain, or for a non-HS medical condition (eg, osteoarthritis, neuropathic pain), the study participant may continue the analgesic. Opioid analgesics (including tramadol) are excluded for any indication.

Notes: (1) Dose should be stable for 14 days prior to Baseline, and is anticipated to remain stable throughout study participation. (2) Dosing PRN is not considered stable, but (for example) taking an NSAID 2 or 3 times per week every week is considered a stable dose.

- If a study participant's pain (HS-related or non-HS-related) worsens after Baseline, the study participant may initiate analgesic therapy at any time and/or per local labeling as follows: For HS-related pain, permitted analgesics are limited to ibuprofen at a dose of up to 800mg orally every 6 hours, not to exceed 3.2g/24 hours; and/or acetaminophen/paracetamol as per local labeling. For non-HS-related pain, initiation of any new analgesic/treatment must not include exclusionary medications (eg, opioids and tramadol), and must be recorded on the eCRF.
 - All analgesic use (start dates, end dates, dose, reason) will be recorded on the eCRF.

- Antibiotic therapy:
 - For study participants entering the study in the antibiotic strata, they should be on a stable dose and regimen of doxycycline, minocycline, or an equivalent systemic tetracycline for 28 days prior to Baseline (Visit 2). The dose and regimen should remain stable throughout study participation, but at least through Week 16. Antibiotics taken on a PRN basis are not considered as a stable dose. After Week 16, participants may receive an antibiotic if required in the judgement of the Investigator. Also see Section 6.5.3.1 for details on systemic antibiotic rescue medication.
 - All antibiotic use (start dates, end dates, dose, reason) will be recorded on the eCRF.

6.5.2 Prohibited concomitant treatments (medications and therapies)

The following concomitant medications and therapies are prohibited during the study (also see Table 6–3):

- All biologic therapy with a potential therapeutic impact on the disease being studied, including those listed in Table 6–3.
- Phototherapy (psoralen and ultraviolet A and/or ultraviolet B) or photochemotherapy.
- Immunomodulatory therapy, including topical or systemic steroids except as noted in Section 6.5.3.1 (Rescue Medications/Lesion Intervention), and Table 6–3.
- Topical and systemic therapies for HS (see Table 6–3).
- Surgical or laser intervention for an HS lesion except as outlined in Section 6.5.3.1 (Rescue Medications/Lesion Intervention).

Table 6–3: Exclusions for prior medications

Drug class	Exclusion criteria
<u>Systemic antibiotics</u>	Used within 2 weeks prior to the Baseline Visit. Note: See exception for permitted, stable doses of antibiotics in Section 6.5.1.
<u>Systemic retinoids</u>	Used within 4 weeks prior to the Baseline Visit
<u>Systemic treatment (non-biologic)</u> <ul style="list-style-type: none"> • Apremilast • Systemic immunosuppressant agents (eg: methotrexate, cyclosporine, azathioprine, thioguanine) • Systemic fumarate • Systemic oral or injectable corticosteroids • Phototherapy and radiotherapy (eg, psoralen and ultraviolet A and/or ultraviolet B) or photoradio/chemotherapy 	Used within 4 weeks prior to the Baseline Visit. Note: See exception for permitted, stable doses of antibiotics in Section 6.5.1.
<u>Anti-tumor necrosis factors (including biosimilars)</u> adalimumab, etanercept, certolizumab, golimumab, infliximab	Used within 12 weeks prior to the Baseline Visit. Note: For etanercept, used within 1 month prior to the Baseline Visit.
<u>Other biologics</u> Abatacept Anakinra Natalizumab Belimumab Tocilizumab Efalizumab Or other biologics approved by regulatory agencies after the protocol is approved	Used within 12 weeks prior to the Baseline Visit Note: for other biologics approved by regulatory agencies after the protocol is approved: Washout of 6 months prior to the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater.
Secukinumab, brodalumab, ixekizumab and other IL-17 inhibitors approved by regulatory agencies after the protocol is approved	Washout of 6 months prior to the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater.
IL-12, IL-23 inhibitors: Ustekinumab Risankizumab Tildrakizumab Guselkumab Or other biologics approved by regulatory agencies after the protocol is approved	Used within 6 months prior to the Baseline Visit Note: for other biologics approved by regulatory agencies after the protocol is approved: Washout of 6 months prior to the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater.

Table 6–3: Exclusions for prior medications

Drug class	Exclusion criteria
<u>Janus kinase inhibitors</u> Tofacitinib Baricitinib Filgotinib Upadacitinib Or other janus kinase inhibitors approved by regulatory agencies after the protocol is approved	Used within 12 weeks of the Baseline Visit or 5 half-lives prior to the Baseline Visit, whichever is greater
Any other systemic HS drug under investigation (or approved after the protocol is approved)	Used within 12 weeks or 5 half-lives prior to the Baseline Visit, whichever is greater
Rituximab	Used within 2 years of the Baseline Visit
Topical drugs for HS (including intralesional corticosteroids, over-the-counter and prescription drugs, as well as disinfectants for skin lesions, eg, chlorhexidine, povidone iodine)	Used within 14 days of the Baseline Visit
Topical corticosteroids (in HS-affected areas) for dermatological use	Used within 14 days of the Baseline Visit. Note: Topical steroids in non-HS affected areas are permitted.
Herbal medications for HS	Used within 14 days of the Baseline Visit
Vaccines	Administration of live (including attenuated) vaccines is not allowed within 8 weeks prior to Baseline, during the conduct of the study, and for 20 weeks after the final dose of IMP (see Exclusion Criteria #23 and #24). Administration of inactivated vaccines is allowed during the study at the discretion of the Investigator. Administration of any other vaccine not mentioned above may be allowed following discussion with the Medical Monitor.
Analgesics	See Section 6.5.1.
Spironolactone	Permitted if indicated for non-HS-related condition (eg, polycystic ovary syndrome); and if study participant meets all other entry criteria. Study participant must be on a stable dose for at least 28 days prior to the Baseline visit.
Metformin	Permitted if indicated for non-HS-related condition (eg, diabetes mellitus); and if study participant meets all other entry criteria. Study participant must be on a stable dose for at least 28 days prior to the Baseline visit.

Table 6–3: Exclusions for prior medications

Drug class	Exclusion criteria
Finasteride and other 5 α -reductase inhibitors	Permitted if indicated for non-HS-related condition (eg, benign prostatic hypertrophy); and if study participant meets all other entry criteria. Study participant must be on a stable dose for at least 28 days prior to the Baseline visit.

HS=hidradenitis suppurativa; IL=interleukin; IMP=investigational medicinal product

6.5.3 Rescue medication

The Sponsor will not supply rescue medication. The following rescue medications may be used:

6.5.3.1 Antibiotic Rescue Medication/Antibiotics

Any systemic antibiotic that is initiated (new antibiotic or change in the dose/type of current antibiotic) on or after Baseline (first day of study drug administration) will be considered rescue medication for both the Initial Treatment Period and Maintenance Treatment Period. If a newly initiated systemic antibiotic (or increase in dose/type of antibiotic) is required during the Initial Treatment Period based on disease flare or other extenuating circumstances, the Investigator should discuss the decision with the Medical Monitor.

6.5.3.2 Rescue Medications/Lesion Intervention

There are no absolute restrictions on the use of rescue medications for study participants whose HS deteriorates during the study. While the objectives of the study should be protected as much as possible through observance of the restrictions detailed above in Section 6.5, the well-being of the study participant will always take priority; study participants should be managed as deemed appropriate by the Investigator.

In the event an acutely painful lesion occurs that requires an immediate intervention, Investigators will have the option to perform interventions. Interventions can include analgesics for a limited period of time (see below), intralesional injections of triamcinolone, and/or incision and drainage of the abscess. Intralesional injections of triamcinolone (up to 20mg across all lesions at a given visit, and using a concentration of no more than 20mg/mL [suspension for injection]) must be consistent with the maximum number of interventions described below and clinical practice. Concomitant use of wound care dressings is permitted; however, options are limited to alginates, hydrocolloids, and hydrogels. Concomitant medications associated with the lesion intervention(s) must be captured in source documents and on the appropriate eCRF.

Any analgesic that is initiated (new analgesic, new class of analgesic, increased dose of an analgesic stable since Baseline, regardless of duration of treatment) after Baseline (first day of study drug administration) will be considered rescue pain medication for the Initial Treatment Period and Maintenance Treatment Period (Weeks 0 to 48).

A total of 2 protocol-allowed interventions are permissible during the Initial Treatment Period (from Baseline Visit to Week 16). Do not include analgesic rescue treatment in the number of protocol-allowed interventions. An intervention can occur on maximally 2 different lesions at the same visit, or on the same lesion at 2 different study visits. The same lesion cannot be treated

2 times at the same visit. If a study participant requires more than 2 interventions within the first 16 weeks of the study, then the study participant should be discontinued from the study.

During the Maintenance Treatment Period (Weeks 16 to 48), a maximum of 2 interventions every 4 weeks are permitted. Do not include analgesic rescue treatment in the number of protocol-allowed interventions. An intervention can occur on 2 different lesions at the same visit or on the same lesion at 2 different study visits. Within each 4-week period, the same type of intervention cannot be used 2 times on the same lesion. If a study participant requires more than 2 interventions within a 4-week period, or has 2 of the same interventions on the same lesion within that period, then the study participant should be discontinued from the study.

6.6 Dose modification

Dose modification is not applicable in this study.

6.7 Criteria for study hold or dosing stoppage

UCB reserves the right to temporarily suspend or prematurely discontinue this study either at a single site, multiple sites, or at all sites at any time for reasons including, but not limited to, safety or ethical issues, inaccurate or incomplete data recording, noncompliance, or unsatisfactory enrollment with respect to quality or quantity. [REDACTED]

If the study is prematurely terminated or suspended, UCB (or its representative) will inform the Investigators/institutions and the regulatory authority(ies) of the termination or suspension and the reason(s) for the termination or suspension, in accordance with applicable regulatory requirement(s). The Institutional Review Board (IRB)/ Independent Ethics Committee (IEC) should also be informed and provided with reason(s) for the termination or suspension by the Sponsor or by the Investigator/institution, as specified by the applicable regulatory requirement(s). In addition, arrangements will be made for the return or destruction of all unused IMP and other material in accordance with UCB procedures for the study.

Detailed procedures for reporting SAEs and other safety events which may meet study hold criteria are provided in Appendix 8.

6.8 Treatment after the end of the study

Study participants who complete HS0004 will have the option of enrolling in a Phase 3, multicenter, extension study (HS0005).

Study participants who elect not to enroll in HS0005 at Week 48 will be scheduled to have the SFU Visit 20 weeks after the final injection of IMP. During the SFU, if study participants' HS deteriorates, the Investigator may consider standard of care for HS treatment after discussion with the Medical Monitor or UCB study physician. Note that the half-life of bimekizumab must be considered in selection of appropriate HS treatments during the SFU period. All concomitant medications and HS interventions administered during the SFU will be recorded on the appropriate eCRF pages.

7 DISCONTINUATION OF STUDY MEDICATION AND PARTICIPANT DISCONTINUATION/WITHDRAWAL

7.1 Discontinuation of study medication

Study medication will be stopped if the study participant develops a medical condition, adverse event (AE), or laboratory abnormality that, in the opinion of the Investigator, compromises the safety of the study participant or his or her ability to continue participation in the study. Study participants who are discontinued from IMP should be encouraged by the Investigator to return for all scheduled visits through Week 48, and the SFU Visit (if the Week 48 Visit is ≥ 20 weeks after the final dose of IMP, the SFU Visit is not required). Any study participant who discontinues IMP but continues in the study should be discussed with the Medical Monitor or UCB study physician.

In all cases the study participant should be followed until the condition has resolved as agreed by the Investigator and the UCB study physician. Investigators should contact the Medical Monitor and/or UCB study physician, in advance whenever possible, to discuss the withdrawal of a study participant.

Study medication will be stopped if the study participant has a confirmed positive coronavirus disease 2019 (COVID-19) test result or a suspected COVID-19 infection. Study medication can be resumed after the participant's recovery from COVID-19, based on the Investigator's clinical judgement. All such cases must be discussed with the Medical Monitor or UCB study physician.

7.1.1 Study participant does not achieve partial response

If a study participant does not achieve a partial response (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at any visit from Week 32 to 46, the Investigator should contact the Medical Monitor to discuss whether the study participant should continue on study.

7.1.2 Potential drug-induced liver injury IMP discontinuation criteria

Study participants with potential drug-induced liver injury (PDILI) must be assessed to determine if IMP must be discontinued. In addition, all concomitant medications and herbal supplements that are not medically necessary should also be discontinued.

7.1.2.1 PDILI Discontinuation Criteria

The PDILI criteria below require immediate discontinuation of IMP for study participants with either of the following (see Section 10.6.2.1):

- ALT or AST $\geq 8 \times \text{ULN}$
- ALT or AST $\geq 3 \times \text{ULN}$ and coexisting total bilirubin $\geq 2 \times \text{ULN}$

Similarly, the PDILI criterion below requires immediate discontinuation of IMP for:

- Study participants with ALT or AST $\geq 3 \times \text{ULN}$ who exhibit temporally associated symptoms of hepatitis or hypersensitivity. Hepatitis symptoms include fatigue, nausea, vomiting, right upper quadrant pain or tenderness. Hypersensitivity symptoms include fever (without clear alternative cause), rash, or eosinophilia (ie, $>5\%$).

If a nondrug-related cause for the symptoms can be confirmed, these study participants may resume IMP administration after discussion with the Medical Monitor and/or UCB study physician, but only when the requirements for rechallenge with IMP as provided in Section 10.6.2.1 are followed.

The PDILI criterion below allows for study participants to continue on IMP at the discretion of the Investigator.

- Study participants with ALT or AST $\geq 5 \times \text{ULN}$ and $< 8 \times \text{ULN}$, total bilirubin $< 2 \times \text{ULN}$, and no eosinophilia (ie, $\leq 5\%$), with no fever, rash, or symptoms of hepatitis (eg, fatigue, nausea, vomiting, right upper quadrant pain or tenderness).

Evaluation of PDILI must be initiated as described in Section 10.6 (Appendix 6) with repeat tests performed in 2 weeks. Upon retest, if ALT or AST values have reduced to $< 5 \times \text{ULN}$, the study participant can continue with the study. However, if ALT or AST remains $\geq 5 \times \text{ULN}$ and $< 8 \times \text{ULN}$ after retest, IMP should be temporarily withheld and study participant should undergo a repeat test in 2 weeks. If ALT or AST values remain $\geq 5 \times \text{ULN}$ even after the second retest, then the study participant should be permanently withdrawn from IMP and should be followed for possible PDILI.

If study participants are unable to comply with the applicable monitoring schedule, IMP must be discontinued immediately.

Investigators should attempt to obtain information on study participants in the case of IMP discontinuation to complete the final evaluation. Study participants with PDILI should not be withdrawn from the study until investigation and monitoring are complete. All results of these evaluations and observations, as well as the reason(s) for IMP discontinuation and study participant withdrawal (if applicable), must be recorded in the source documents. The eCRF must document the primary reason for IMP discontinuation.

7.1.2.1.1 Follow-up evaluation

Potential drug-induced liver injury events require follow-up monitoring as described in Table 10–1. Monitoring should continue until liver chemistry values normalize, stabilize, or return to Baseline. Determination of stabilization is at the discretion of the Investigator in consultation with the hepatologist (as applicable) and UCB study physician, as needed.

7.1.3 Treatment interruptions

If a study participant is found to be persistently noncompliant (for example, missing 2 or more of the doses in the Initial Treatment Period or 3 or more doses during Maintenance Treatment Period) the Sponsor, in conjunction with the Investigator, will make a decision as to whether the study participant should be withdrawn from the study.

Note: Doses of IMP that were missed due to a reasonable interfering AE that does not allow administration of IMP due to safety reasons will not be considered for the evaluation of study participant discontinuation. Evaluation of the reasonability of the AE should be discussed immediately with the Medical Monitor.

Any participant who develops a clinically important infection or recurrent infections not responsive to standard therapy during the study must discontinue IMP until resolution of the

infection. The Investigator should use clinical judgement in deciding whether the participant should restart IMP and contact the Medical Monitor and UCB study physician to confirm the participant's suitability for continued participation in the study.

7.2 Participant discontinuation/withdrawal from the study

Note: For female study participants, please see Section 8.3.5 for pregnancy that occurs during the study as evidenced by a positive pregnancy test.

Study participants are free to withdraw from the study at any time, without prejudice to their continued care.

A study participant may withdraw from the study at any time at his or her own request, or may be withdrawn at any time at the discretion of the Investigator for safety, behavioral, compliance, or administrative reasons.

If the study participant withdraws consent for disclosure of future information, the Sponsor may retain and continue to use any data collected before such a withdrawal of consent.

If a study participant withdraws from the study, he or she may request destruction of any samples taken and not tested, and the Investigator must document this in the site study records.

See the Schedule of Activities (Table 1–1) for data to be collected at the time of study discontinuation and follow-up and for any further evaluations that need to be completed.

If a study participant does not achieve a partial response (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at any visit from Week 32 to 46, the Investigator should contact the Medical Monitor to discuss whether the study participant should continue on study.

Study participants will be withdrawn from the study, after being encouraged to complete the Premature End of Treatment (PEOT) and the SFU Visit if either of the following events occur:

1. Study participant withdraws his or her consent.
2. The Sponsor or a regulatory agency requests withdrawal of the study participant.

Study participants should be withdrawn from IMP and encouraged by the Investigator to return for all scheduled visits through Week 48, and the SFU Visit (if the Week 48 Visit is ≥ 20 weeks after the final dose of IMP, the SFU Visit is not required) if any of the following events occur:

1. Study participant develops an illness that would interfere with his or her continued participation.
2. Study participant is noncompliant with the study procedures or medications in the opinion of the Investigator.
3. Study participant takes prohibited concomitant medications as defined in this protocol (Section 6.5.2) that may present a risk to the safety of the participant or the integrity of the study data, in the opinion of the Investigator and/or the Medical Monitor and/or UCB study physician.
4. Study participant requires more than the number of protocol-allowed lesion interventions (see Section 6.5.3.2).

5. Study participant has a clinical laboratory value meeting any of the following criteria:

- a. Hepatotoxicity as described in Section 7.1.2.
- b. A laboratory value meeting any of the following criteria:
 - Absolute neutrophil count $<1.0 \times 10^3/\mu\text{L}$
 - Absolute lymphocyte count $<200 \text{ cells}/\mu\text{L}$

Study participants may remain on IMP if the result is transient. A retest is required within 1 to 2 weeks at a scheduled or unscheduled visit. If the repeat absolute neutrophil count or absolute lymphocyte count is still below the allowable values, the participant must be discontinued from the IMP. If the repeat absolute neutrophil count or absolute lymphocyte count is above the allowable values, the participant may continue to receive IMP.

6. There is confirmation of a pregnancy during the study, as evidenced by a positive pregnancy test (see Section 8.3.5 for more information regarding pregnancies).
7. A study participant considered as having either a suspected new LTBI or who develops active TB or an NTM infection during the study (including but not limited to, conversion demonstrated by interferon gamma release assay [IGRA] or other diagnostic means) must be immediately discontinued from IMP.
 - The study participant must immediately be permanently withdrawn from the study if further examinations result in a diagnosis of the following:
 - active TB or
 - an NTM infection, or
 - latent TB infection and study participant does not initiate TB prophylactic therapy, prematurely discontinues TB prophylactic therapy, or, in the opinion of the Investigator or Sponsor, is noncompliant with TB prophylactic therapy.

The PEOT Visit must be scheduled as soon as possible, but not later than the next regular visit.

- If the study participant is diagnosed with LTBI during the study and desires to continue in the study, he or she must immediately discontinue IMP and start TB prophylactic therapy. After at least 4 weeks of TB prophylaxis, the IMP can be restarted after discussion with the UCB study physician regarding results of laboratory assessments, physical examination, and TB questionnaire. The full course of TB prophylaxis treatment will be completed during the study.

Confirmed active TB is an SAE and must be captured on an SAE Report Form and provided to the Sponsor in accordance with SAE reporting requirements. As with all SAEs, periodic follow-up reports should be completed as per protocol requirements until such time as the TB infection resolves.

Additional information on TB policies is provided in Section 8.2.6.

8. Study participants with newly diagnosed IBD or with IBD flares during the study must:

- Be referred, as appropriate, to a healthcare professional treating IBD, such as a gastroenterologist
- Discontinue IMP and be followed-up until resolution of active IBD symptoms

If IBD flares increase in severity or frequency during the study, the Investigator should use clinical judgement in deciding whether the study participant should continue on IMP and contact the Medical Monitor and UCB study physician to confirm the study participant's suitability for continued participation in the study.

- 9 Study participants must be referred immediately to a mental healthcare professional and may be withdrawn from the study based upon the Investigator's judgment of benefit/risk for:
- Active suicidal ideation as indicated by a positive response ("Yes") to Question 4 of the "Since Last Visit" version of the eC-SSRS
 - Moderately severe major depression as indicated by a PHQ-9 score of 15 to 19 if this represents an increase of at least 3 points compared to the last visit

The mental health consultation must be recorded in the study participant's source documentation.

10. Study participants must be referred immediately to a mental healthcare professional and must be withdrawn from the study in case of:

- Active suicidal ideation as indicated by a positive response ("Yes") to Question 5 of the "Since Last Visit" version of the eC-SSRS.
- Any suicidal behavior since last visit.
- Severe major depression as indicated by a PHQ-9 score ≥ 20 .

The mental health consultation must be recorded in the study participant's source documentation.

Investigators should contact the Medical Monitor in advance, whenever possible, to discuss the withdrawal of a study participant from IMP or from the study.

Study participants withdrawing from the study who are not continuing for all scheduled visits through Week 48, will undergo the PEOT Visit and the SFU Visit 20 weeks after their final dose of IMP, as applicable.

The eCRF must document the primary reason for withdrawal or discontinuation.

Investigators should contact the Medical Monitor and/or UCB study physician, whenever possible, to discuss the withdrawal of a study participant in advance.

Withdrawn participants will not be replaced.

7.3 Lost to follow up

A study participant will be considered lost to follow-up if he or she repeatedly fails to return for scheduled visits and is unable to be contacted by the study site.

The following actions must be taken if a study participant fails to return to the clinic for a required study visit:

- The site must attempt to contact the study participant and reschedule the missed visit as soon as possible and counsel the study participant on the importance of maintaining the assigned visit schedule and ascertain whether or not the study participant wishes to and/or should continue in the study.
- Before a study participant is deemed lost to follow up, the Investigator or designee must make every effort to regain contact with the study participant (at least 1 phone call and 1 written message to the participant), and document his or her effort (date and summary of the phone call and copy of the written message in the source documents), to complete the final evaluation (PEOT and SFU, as applicable). All results of these evaluations and observations, together with a narrative description of the reason(s) for removing the study participant, must be recorded in the source documents. The CRF must document the primary reason for withdrawal.

Should the study participant continue to be unreachable, he or she will be considered to have withdrawn from the study with a primary reason of lost to follow-up documented in the eCRF.

8 STUDY ASSESSMENTS AND PROCEDURES

Study procedures and their timing are summarized in the Schedule of Activities ([Table 1-1](#)).

Some study-specific assessments may be conducted remotely according to the study protocol during a pandemic or other exceptional circumstance (eg, hurricanes) due to the need to implement safety measures and guidance from regulatory authorities. In such a situation, which may be accompanied by local or global containment or other measures, sites may need to prohibit access to study participants and study-related personnel. Study participants' visits to the study site may be replaced by contingency measures. These measures are primarily established to ensure the safety of study participants during the course of the study and to maintain the study participants' treatment schedules, if the Investigator considers it appropriate. These measures include, but are not limited to, virtual visits or home-nursing visits replacing site visits, eg, telemedicine contacts or home-nursing visits when treatment and/or blood sampling is scheduled. Any study specific assessments not conducted due to such circumstances must be recorded appropriately in the source documents and eCRF. If it is related to the COVID-19 pandemic, then it must be captured in the COVID-19 impact eCRF page. The contingency measures are described in a contingency plan that will be maintained by UCB for the respective study. The contingency measures are shared with the Investigator and the respective study-related personnel as soon as there are indications that it is necessary to implement any of the measures.

Protocol waivers or exemptions are not allowed.

Immediate safety concerns should be discussed with the Medical Monitor and/or UCB study physician immediately upon occurrence or awareness to determine if the study participant should continue or discontinue study treatment.

Adherence to the study design requirements, including those specified in the Schedule of Activities, is essential and required for study conduct.

All screening evaluations must be completed and reviewed to confirm that potential study participants meet all eligibility criteria. The Investigator will maintain a screening log to record details of all study participants screened and to confirm eligibility or record reasons for screening failure, as applicable.

Procedures conducted as part of the study participant's routine clinical management (eg, blood count) and obtained before signing of the ICF may not be utilized for screening or Baseline purposes.

The maximum amount of blood collected from each study participant over the duration of the study, including any extra assessments that may be required, will not exceed the usual volume of blood taken for a blood donation. Repeat or unscheduled samples may be taken for safety reasons or for technical issues with the samples.

8.1 Efficacy assessments

The timing for all assessments described below is specified in [Table 1–1](#).

8.1.1 Lesion count

The lesion count is defined as an assessment of all the various skin “appearances” that are termed “lesions” in HS study participants. The lesion count will include the following:

- Abscesses (circumscribed collection of purulent exudate frequently associated with swelling and other signs of inflammation, such as fluctuance, tenderness, and pain)
- Draining tunnels (fistulas/sinus tracts) (pathologic passageway connecting to the skin surface from dermis or sc tissue/pathologic passageway that develops into a channel to the skin surface that drains serous or purulent fluid, either spontaneously or by gentle palpation)
- Non-draining tunnels (fistulas/sinus tracts) (pathologic passageway connecting to the skin surface from dermis or sc tissue/pathologic passageway that develops into a channel to the skin surface that does not drain serous or purulent fluid)
- Noninflammatory nodules (nontender or minimally tender, nonerythematous nodules)
- Inflammatory nodules (a tender, erythematous, well-defined nodule. The lesion has no evidence of fluctuance. A pyogenic granuloma lesion is considered an inflammatory nodule; a papule or pustule is not considered an inflammatory nodule)
- Scars of HS lesions (enlargement or overgrowth of a scar so that it extends above the surrounding skin surface)

The data collected from the lesion count will be used for the derivation of study variables including, but not limited to HiSCR₂₅ (a 25% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count), HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀ (a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count), HS Physician's Global Assessment, AN count, and International HS Severity score system (IHS4).

8.1.1.1 Hidradenitis Suppurativa Clinical Response (HiSCR)

The HiSCR; defined as at least a 50% reduction from Baseline in the total AN count, with no increase from Baseline in abscess or draining tunnel count; was developed to address issues with

available HS scoring systems. It is a validated endpoint that is responsive to improvement in disease activity, simplifies the scoring process, and increases the sensitivity to detect HS-specific lesions (Kimball, 2014; Kimball, 2016b). HiSCR has been labeled HiSCR₅₀ in this protocol.

The HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ also will be evaluated in this study. These measures of clinical response differ from HiSCR₅₀ only in the percent decrease in AN count from Baseline.

The HiSCR_{xx} is derived by the statistical programming group based on Investigator documentation of lesion count, and does not require calculation on the part of the Investigator from the lesion count.

8.1.1.2 Hidradenitis Suppurativa Physician's Global Assessment

The HS Physician's Global Assessment is a validated 6-point scale that is used to measure improvement in inflammatory nodules, abscesses, and draining tunnels (Kimball, 2012; Zouboulis, 2015). The HS Physician's Global Assessment scale is defined by the following:

- Clear: No inflammatory or noninflammatory nodules
- Minimal: Only the presence of noninflammatory nodules
- Mild: ≥ 1 and ≤ 4 inflammatory nodules or 1 abscess or draining tunnel and no inflammatory nodules
- Moderate: ≥ 5 inflammatory nodules or 1 abscess or draining tunnel and 1 or more inflammatory nodules or 2 to 5 abscesses or draining tunnels and ≤ 10 inflammatory nodules
- Severe: 2 to 5 abscesses or draining tunnels and > 10 inflammatory nodules
- Very severe: > 5 abscesses or draining tunnels

This assessment (clear, minimal, mild, moderate, severe, or very severe) is derived based on totals across all affected body regions by the statistical programming group and does not require calculation on the part of the Investigator.

8.1.1.3 International Hidradenitis Suppurativa Severity Score System (IHS4)

The IHS4 is a validated tool to dynamically assess HS severity to be used both in real-life and the clinical trials setting (Zouboulis, 2017). The IHS4 achieved consensus among European HS Foundation members. This IHS4 score is calculated as follows: (number of nodules multiplied by 1) + (number of abscesses multiplied by 2) + [number of draining tunnels (fistulae/sinuses) multiplied by 4]. A score of 3 or less signifies mild HS, a score of 4–10 signifies moderate HS and a score of 11 or higher signifies severe HS.

The determination of IHS4 requires counting the nodules, abscesses and draining tunnels/sinus tracts.

The IHS4 score is derived by the statistical programming group, and does not require calculation on the part of the Investigator.

8.1.2 Partial response

A partial response is defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at a particular timepoint.

The partial response is derived by the statistical programming group, and does not require calculation on the part of the Investigator.

8.1.3 High-sensitivity C-reactive protein (hs-CRP)

Blood will be collected for measurement of hs-CRP. The hs-CRP data will not be sent to any blinded study personnel to protect the blinded nature of the treatment assignments and response.

8.1.4 Patient-reported outcomes

The patient-reported outcome (PRO) instruments should be completed by the study participants themselves in a quiet place. The PRO instruments to be completed at the study site, should be completed prior to all other protocol-specified assessments at each visit (including dosing on dosing days).

8.1.4.1 HS symptom measures of skin pain, smell or odor, drainage or oozing from HS lesions, and itch

8.1.4.1.1 HS symptom daily diary (HSSDD)

The 5 items on the HS Symptom Daily Diary (HSSDD) assesses patients' perception of the core symptoms of HS experienced in the past 24 hours: pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point numeric rating scale (NRS). Two items assess skin pain (ie, worst skin pain and average skin pain). The remaining 3 items assess smell or odor, itch at its worst, and amount of drainage or oozing from HS lesions.

The HSSDD will be completed daily by the study participant, at the end of the day on an electronic hand-held device from the start of Screening through the Week 16 visit.

8.1.4.1.2 HS symptom questionnaire (HSSQ)

The 4 items on the HS Symptom Questionnaire (HSSQ) assesses patients' perception of the core symptoms of HS experienced in the past 7 days - pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point NRS. The HSSQ will be completed on an electronic device during study visits (ie, Baseline and Weeks 16-48/PEOT).

8.1.4.2 Patient Global Impression of HS Severity (PGI-S-HS) and Change in HS Severity (PGI-C-HS)

The Patient Global Impression of HS Severity (PGI-S-HS) is a single item to assess study participants' perceptions of the overall severity of HS over the past 7 days (none, mild, moderate, severe, very severe). The Patient Global Impression of Change in HS Severity (PGI-C-HS) is a single item to assess study participants' perception of the change in HS since they started taking the study medication (much better, a little better, no change, a little worse, much worse). Data collected using the PGI-S-HS and PGI-C-HS will be used as anchors for interpreting change scores on the Hidradenitis Suppurativa Quality of Life (HiSQOL).

8.1.4.3 Patient Global Impression of Severity of Skin Pain (PGI-S-SP) and Change in Severity of Skin Pain (PGI-C-SP)

The Patient Global Impression of Severity of Skin Pain (PGI-S-SP) is a single item to assess study participants' perceptions of the severity of their skin pain from their HS lesions, over the past 7 days (none, mild, moderate, severe, very severe). The Patient Global Impression of Change in Severity of Skin Pain (PGI-C-SP) is a single item to assess study participants'

perceptions of change in their skin pain from their HS lesions, since they started taking the study medication (much better, a little better, no change, a little worse, much worse). PGI-S-SP and PGI-C-SP will be used to evaluate outcomes related to Skin Pain.

8.1.4.4 Hidradenitis Suppurativa Quality of Life Questionnaire (HiSQOL)

The 17 item HiSQOL questionnaire has a recall period of 7 days. The HiSQOL includes 3 subscales: symptom status, psychosocial impact, and impact on physical activities.

8.1.4.5 Dermatology Life Quality Index (DLQI)

The DLQI is a questionnaire designed for use in adult participants with inflammatory skin diseases and has been used in patients with HS (Finlay, 1998; Esmann, 2010; Basra, 2012). The DLQI is a skin disease-specific questionnaire aimed at the evaluation of how symptoms and treatment affect participants' health-related QOL. This instrument asks participants about symptoms and feelings, daily activities, leisure, work and school, personal relationships and treatment. The DLQI total score ranges from 0 to 30 with higher scores indicating lower health related QOL. In other dermatological/skin conditions, a 4-point change in the DLQI total score (DLQI response) has been reported to be meaningful for the participant (within participant minimal important difference); while a DLQI total absolute score of 0 or 1 indicates no or small impact of the disease on health related QOL.

8.1.4.6 Euro-Quality of Life 5-Dimensions, 3 levels

The EQ-5D-3L comprises 5 questions on mobility, self-care, pain, usual activities, and psychological status with 3 possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). In addition, the questionnaire includes a visual analogue scale to indicate the general health status, with 100 indicating the best health status.

8.1.4.7 Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem

The Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem (WPAI-SHP) V2.0 is a patient-reported questionnaire that assesses study participant's employment status, work absenteeism, work impairment while working (presenteeism), overall work, and daily activity impairment attributable to a specific health problem (Reilly, 1993). It has been used in several clinical studies of biologic therapy in participants with plaque PSO (Kimball, 2012; Vender, 2012).

Five out of 6 items of the WPAI-SHP are regrouped into the 4 dimensions, with scores expressed as percentage, where higher numbers indicate greater impairment and less productivity, ie, worse outcomes, as described in the WPAI-SHP scoring rules.

8.1.4.8 Treatment Satisfaction Questionnaire for Medication

The Treatment Satisfaction Questionnaire for Medication (TSQM-9) is an abbreviated 9-item version of the TSQM, excluding the side effects of medication domain. The domains included in the TSQM-9 include effectiveness (3 items), convenience (3 items) and global satisfaction (3 items). The TSQM-9 domain scores range from 0 to 100 with higher scores representing higher satisfaction (Bharmal, 2009).

8.1.5 Hurley Stage

The Hurley Stage is a severity classification for HS that was developed in 1989 and is widely used for the determination of the severity of HS (Hurley, 1989).

The Hurley Stage is defined by the following criteria:

- Stage I: Abscess formation, single or multiple, without sinus tracts and cicatrization (scarring).
- Stage II: Recurrent abscesses with tract formation and cicatrization, single or multiple, widely separated lesions.
- Stage III: Diffuse or near-diffuse involvement, or multiple interconnected tracts and abscesses across the entire area.

Hurley Stage is assigned to a given anatomic region. The overall worst Hurley Stage (ie, the highest Hurley Stage across all anatomic regions) for a given study participant at a given visit is then the study participant-level Hurley Stage. This study participant-level Hurley Stage is important for baseline stratification and disease severity assessment.

Hurley stage is included as a stratification factor for randomization.

8.2 Safety assessments

Planned timepoints for all safety assessments are provided in the Schedule of Activities (Table 1–1).

8.2.1 Physical examination

A complete physical examination will include, at a minimum, general appearance; ear, nose, and throat; eyes, hair, and skin; and assessments of the cardiovascular, respiratory, gastrointestinal musculoskeletal, and hepatic, neurological (including limb reflexes) systems, and mental status. Each physical examination also includes evaluation of signs and symptoms of active TB and risk for exposure to TB (Section 8.2.6).

Height and weight will also be measured and recorded. The same scale for measuring body weight should be utilized throughout the study where possible.

Investigators should pay special attention to clinical signs related to previous serious illnesses.

Clinically relevant findings or worsening of previous findings since the physical examination at the Screening Visit will be recorded as AEs.

8.2.2 Vital signs

Vital signs will be measured in a sitting position after 5 minutes rest and will include body temperature (oral, axillary, otic or noncontact forehead), systolic and diastolic blood pressure, and pulse. Vital signs are to be measured prior to blood sampling, and prior to dosing, where applicable.

Vital signs will consist of single pulse and blood pressure measurements.

8.2.3 Electrocardiograms

A single 12-lead ECG will be obtained using an ECG machine that automatically calculates the heart rate and measures PR, QRS, QT, and QT corrected for heart rate intervals.

All ECG recordings should be taken with the study participant resting in the supine position for at least 10 minutes before the recording and prior to taking blood samples or dosing.

ECG machines will be provided to study centers, and ECGs will be read by a central ECG laboratory. Full details of ECG recordings will be provided in the ECG Manual.

8.2.4 Clinical safety laboratory assessments

See Appendix 2 (Section 10.2) for the list of clinical laboratory tests to be performed.

The Investigator must review the laboratory report, document this review, and record any clinically relevant changes occurring during the study in the AE section of the CRF. The laboratory reports must be filed with the source documents. Clinically significant abnormal laboratory findings are those which are not associated with the underlying disease, unless judged by the Investigator to be more severe than expected for the study participant's condition.

All laboratory tests with values considered clinically significantly abnormal during participation in the study or within 20 weeks after the last dose of IMP should be repeated until the values return to normal or Baseline or are no longer considered clinically significant by the Investigator or Medical Monitor.

If such values do not return to normal/Baseline within a period of time judged reasonable by the Investigator, the etiology should be identified and the Sponsor notified.

All protocol-required laboratory assessments, as defined in Appendix 2, must be conducted in accordance with the laboratory manual and the Schedule of Activities.

If laboratory values from non-protocol specified laboratory assessments performed at the institution's local laboratory require a change in participant management or are considered clinically significant by the Investigator (eg, SAE or AE or dose modification), then the results must be recorded in the CRF.

8.2.5 Depression and suicidal risk monitoring

8.2.5.1 PHQ-9

The PHQ-9 is a multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression. The PHQ-9 scores for depression range from 0 to 27 with higher scores indicating worse state. A score of 5 to 9 is considered to be minimal symptoms of depression. A score of 10 to 14 is considered minor depression, dysthymia, or mild major depression. A score of 15 to 19 is considered to indicate moderately severe major depression, and a score ≥ 20 is considered to be severe major depression.

Refer to Section 7.2 for PHQ-9-related withdrawal criteria.

8.2.5.2 eC-SSRS

The eC-SSRS is a standardized and validated instrument developed for the assessment of the severity and frequency of suicidal ideation and behavior (Mundt, 2010; Posner, 2011). Study participants respond to standardized clinical questions that are presented in a uniform fashion.

The eC-SSRS defines 5 subtypes of suicidal ideation and behavior in addition to self-injurious behavior with no suicidal intent. The eC-SSRS takes approximately 3 to 10 minutes to complete.

Refer to Section 7.2 for eC-SSRS-related withdrawal criteria.

8.2.6 Assessment and management of tuberculosis and tuberculosis risk factors

All participants will be assessed for TB through physical examination for signs and symptoms of TB, laboratory testing (Section 8.2.4), chest x-ray (CXR) (Section 8.2.6.3.2), and TB questionnaire (Section 8.2.6.3.3).

8.2.6.1 Assessments at Screening

At Screening, all participants will have an IGRA test (QuantiFERON Gold Plus TB test is recommended), a CXR (unless already performed within 2 months of Screening; a computed axial tomography (CAT) scan of the chest at Screening or within 2 months prior to Screening is acceptable, if available), and examination for signs and symptoms of TB. In addition, the Investigator or designee will complete a TB questionnaire directed at the participants potential exposure to TB and symptoms of TB.

Study participants diagnosed with active TB during Screening will be excluded from the study.

Study participants with LTBI diagnosed during Screening must have completed a full course of prophylaxis prior to IMP dosing and can be rescreened after completion of a full course of prophylaxis plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline. (See also Section 8.2.6.3.5.)

8.2.6.2 Definitions

Study participants with known active TB disease, at high risk of acquiring TB infection, or with untreated LTBI (ie, pending anti-TB prophylactic course) or current or history of NTM infection are excluded from the study.

a. Known TB infection whether present or past is defined as:

- Active TB disease or clinical signs and symptoms strongly suggestive of TB (pulmonary or extra pulmonary).
- History of active TB disease involving any organ system or findings in other organ systems consistent with TB, unless adequately treated and proven to be fully recovered upon consult with a TB specialist.
- Any evidence by radiography or other imaging modalities consistent with previously active TB disease that is not reported in the study participant's medical history.

b. High risk of acquiring TB infection is defined as:

- Known close exposure (eg, sleeping in the same room) to another person with active TB infection within 3 months prior to Screening.
- Time spent within 3 months prior to screening in a healthcare delivery setting or institution where individuals infected with TB are housed or where the risk of transmission of infection is high.

- c. Latent TB infection is defined as an infection by *Mycobacterium tuberculosis* with:
- A positive IGRA (or 2 indeterminate IGRAs) AND
 - Chest imaging (or other imaging) negative for TB infection, AND
 - Absence of signs, symptoms (eg, evidence of organ-specific involvement), or physical findings suggestive of TB infection.
- d. Pulmonary NTM infection is defined as a group of lung or extrapulmonary infections caused by mycobacteria different from *M. tuberculosis* infections.

8.2.6.3 Assessment and reporting of TB and TB risk factors during the study

8.2.6.3.1 Physical examination

The Investigator should consider all potential sites of infection when assessing for TB during the physical examination, and other evaluations, and based on the study participant's medical or social history.

The most common primary focus of TB is the lung. Other sites may include gastrointestinal system, bone/joints, lymph glands and meninges, etc. However, in immune-compromised patients, study participants, and/or patients treated with biologics, especially tumor necrosis factors inhibitors, extra-pulmonary manifestations of TB is common compared to normal population.

Some common symptoms that the study participant may present are dependent on the primary focus of infection and may include cough, blood in sputum, night sweats, lymphadenitis, joint pain/swelling, spinal deformity, headache/confusion, abdominal pain (mimicking inflammatory bowel disease), etc. Unusual presentations should always be considered.

8.2.6.3.2 Chest x-ray for tuberculosis

Chest radiographic imaging is performed at screening and results must be available at baseline before first drug administration unless a CXR or CAT scan is available from 2 months prior to screening.

Additional CXR or other imaging test should be performed when positive signs and symptoms indicate pulmonary infection, including potential TB infection, or when close exposure to persons with TB is documented.

8.2.6.3.3 Tuberculosis questionnaire

A questionnaire entitled "Evaluation of Signs and Symptoms of Tuberculosis" has been developed by UCB (document mod-000582) to help in identifying TB risk factors in study participants; it is administered by the Investigator or their designee. For the purpose of case reporting, this questionnaire also ensures appropriate follow-up with reporters when a case of either latent TB or active TB is diagnosed. Moreover, it ensures proactive and appropriate follow-up with Investigators and study participants on treatment course.

8.2.6.3.4 IGRA Test Conversion

The IGRA is a whole-blood testing methodology for diagnosing *M. tuberculosis* infection. It has become the gold standard, but does not help in differentiating LTBI from active tuberculosis disease.

Tuberculosis test conversion is defined as a positive or indeterminate (and confirmed indeterminate on repeat) IGRA result for the current test when previous IGRA test results were negative. All study participants with positive or indeterminate IGRA test results must immediately stop IMP administration. In case of a IGRA test conversion, the study participant must be considered as having either a suspected new latent or an active TB infection and be promptly referred to an appropriate specialist (eg, pulmonologist, infectious disease specialist) for further evaluation. Additional assessments (eg, blood tests or IGRA, CXRs, or other imaging) should be performed where medically relevant and documented. Such conversions should be reported as AEs as described in the protocol. The AE term would need to be updated with final diagnosis once available.

8.2.6.3.5 Latent TB

In case the evaluation by the appropriate specialist diagnoses a new LTBI, a TB prophylactic therapy in accordance with applicable clinical guidelines should be immediately initiated.

Study participants who initiate treatment for LTBI during the Screening period must repeat initial screening laboratory parameters, all physical examinations, and questionnaires prior to randomization in the study, and must continue the full course of TB prophylactic therapy. Participants can be rescreened after completion of a full course of prophylaxis. Eligible study participants can be included after a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline.

Study participants who initiate treatment for LTBI during the study must repeat study assessments after TB prophylactic therapy has been received for at least 4 weeks. The Investigator and Medical Monitor and/or UCB study physician will decide which investigations (safety laboratory parameters, physical exams and questionnaires) need to be performed after required LTBI prophylaxis period and before the IMP will be restarted.

The IMP can be restarted no sooner than 4 weeks after the start of TB prophylactic therapy if it is deemed likely that the TB prophylactic therapy will be continued to full completion. If no TB prophylactic therapy is initiated for the newly diagnosed LTBI, the study participant must permanently stop IMP and be withdrawn from the study. Every related action should be discussed in advance with the Medical Monitor.

Study participants who prematurely discontinue treatment for LTBI or who, in the opinion of the Investigator or Sponsor, are noncompliant with anti-TB therapy must discontinue further intake of IMP and be immediately withdrawn. Once withdrawn from study treatment, study participants should return for the PEOT visit, complete all assessments, and complete the SFU visit. LTBI must be reported as an AE. Follow-up reports should be completed as per protocol requirement until such time as the LTBI resolves.

8.2.6.3.6 Active TB or non-tuberculosis mycobacterium infection

Study participants who develop active TB or NTM infection during the study must be withdrawn from the study. The study participant must be immediately permanently discontinued from study medication and a PEOT visit must be scheduled as soon as possible, but no later than the next scheduled visit. The study participant should be encouraged to keep the SFU visit as specified by the protocol. Treatment for active TB or NTMB should be started immediately.

Confirmed active TB is always considered an SAE. UCB's process requires that these must be captured on an SAE report form and provided to the Sponsor in accordance with SAE reporting requirements. Follow-up reports should be completed as per protocol requirement until such time as the TB infection resolves.

8.2.6.3.7 Tuberculosis management of LTBI, active TB, or other NTB infection identified during study

During the study, study participants who develop evidence of LTBI, active TB or NTB infection must immediately stop further administration of IMP and will be referred to a TB specialist (pulmonologist or infectious disease specialist) for further evaluation. Study participants diagnosed with active TB or LTBI should receive appropriate TB or prophylaxis therapy. If a TB specialist excludes active TB, the study participant can restart the IMP no earlier than 4 weeks after the start of an appropriate TB prophylactic therapy. The study participant should be transferred to the care of his or her physician and managed according to the standard of care.

Study participants identified as having active TB during the study must be withdrawn and scheduled to return for the PEOT Visit as soon as possible but no later than the next scheduled study visit and complete all PEOT Visit assessments. The study participant should be encouraged to complete an SFU Visit after the last dose of study medication.

If infection with NTM is identified during the study, the same procedure as for active TB acquired during the study and compliant TB treatment shall be followed.

Study participant eligibility, retesting requirements, and treatment requirements are shown in [Figure 8-1](#) (screening) and [Figure 8-2](#) (during the study). Additional details on TB detection and management are provided in the UCB TB Detection Procedure Guideline.

Figure 8-1: Decision tree for IGRA TB results at Screening

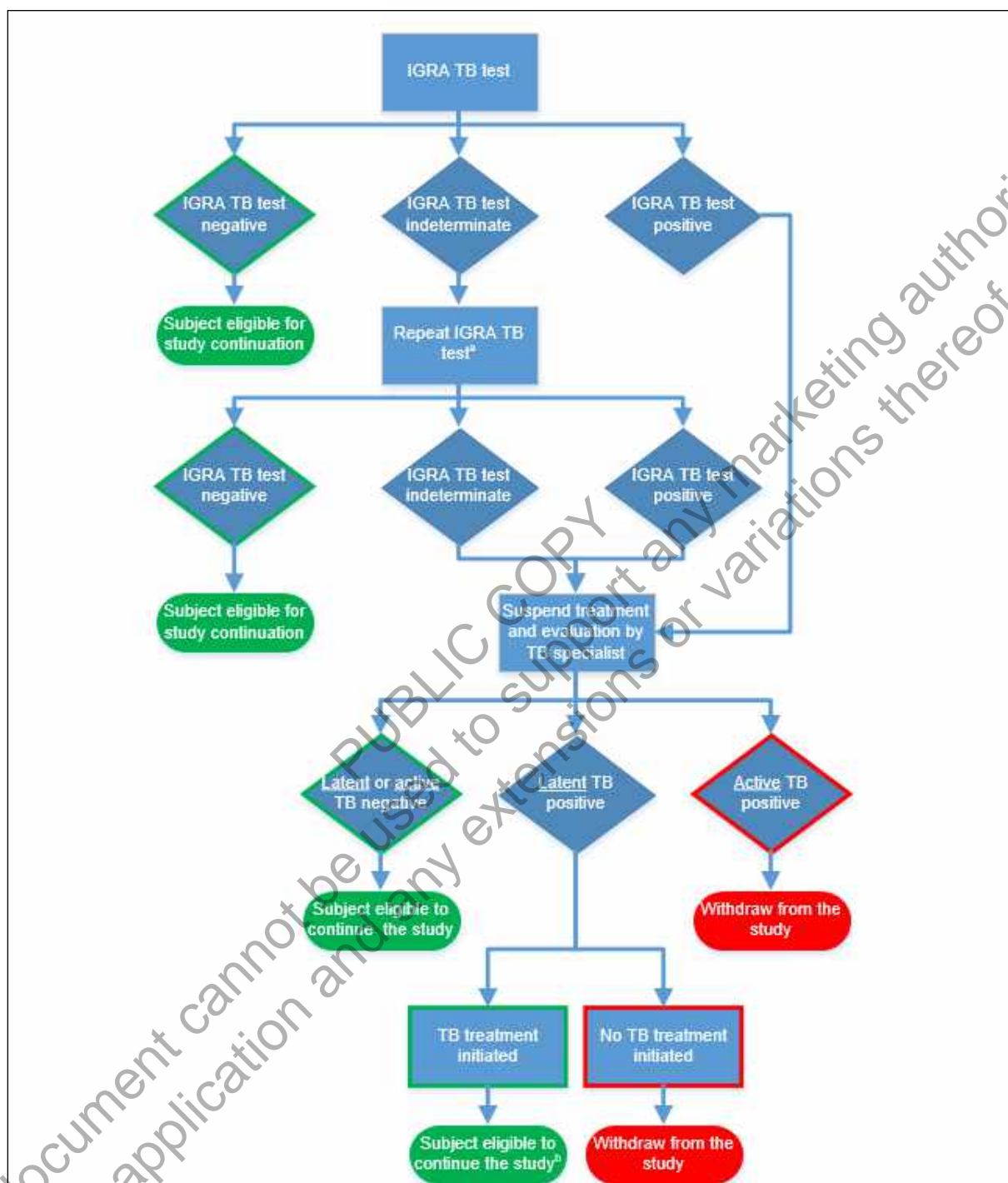


IGRA=interferon gamma release assay; IMP=investigational medicinal product; LTBI=latent tuberculosis infection; TB=tuberculosis

^a IGRA retest must be done during the protocol-defined Screening window

^b Study participants with LTBI diagnosed during Screening must have completed a course of prophylaxis prior to IMP dosing. Participants can be rescreened after completion of a course of prophylaxis plus a wash-out of least 5 half-lives of the prophylactic medication(s) prior to Baseline.

Figure 8-2: Decision tree for IGRA TB results during a study



ASAP=as soon as possible; IGRA=interferon-gamma release assay; IMP=investigational medicinal product; LTBI=latent tuberculosis infection; TB=tuberculosis

^a IGRA retest must be done ASAP and prior to the next IMP dose

^b Study participants with LTBI diagnosed during the study may continue the study only after they have completed at least 4 weeks of appropriate prophylactic therapy and thereafter, will continue and complete the entire regimen.

8.3 Adverse events and serious adverse events

AE will be reported by the study participant (or, when appropriate, by a caregiver, surrogate, or the study participant's legally authorized representative).

The Investigator and any designees are responsible for detecting, documenting, and recording events that meet the definition of an AE or SAE and remain responsible for following up AEs that are serious, considered related to the study treatment or study procedures, or that caused the study participant to discontinue the study treatment or the study (see Section 7).

Confirmed and suspected cases of COVID-19 infection will be recorded as AEs (or SAEs, as required).

8.3.1 Time period and frequency for collecting AE and SAE information

All AEs and SAEs will be collected from the signing of the ICF until the SFU visit (except for those study participants who enroll in extension study HS0005) or until the first dose administration in extension study HS0005 (for study participants enrolling in HS0005).

In order to ensure complete safety data collection, all AEs occurring during the study (ie, after the signing of the ICF), including any pretreatment and posttreatment periods required by the protocol, must be reported in the CRF even if no IMP was taken but specific study procedures were conducted. This includes all AEs not present prior to the initial visit and all AEs that recurred or worsened after the initial visit.

All SAEs will be recorded and reported to the Sponsor or designee within 24 hours, as indicated in Appendix 3. The Investigator will submit any updated SAE data to the Sponsor within 24 hours of it being available.

The Investigator is specifically requested to collect and report to UCB (or its representative) any SAEs (even if the Investigator is certain that they are in no way associated with the IMP), up to 20 weeks from the last dose of IMP for each study participant, and to also inform study participants of the need to inform the Investigator of any SAE within this period. Serious AEs that the Investigator thinks may be associated with the IMP must be reported to UCB regardless of the time between the event and the end of the study.

The method of recording, evaluating, and assessing causality of AE and SAE and the procedures for completing and transmitting SAE reports are provided in Section 10.3.

8.3.2 Method of detecting AEs and SAEs

Care will be taken not to introduce bias when detecting AEs and/or SAEs. Open-ended and non-leading verbal questioning of the participant is the preferred method to inquire about AE occurrences.

8.3.3 Follow-up of AEs and SAEs

After the initial AE/SAE report, the Investigator is required to proactively follow each study participant at subsequent visits/contacts. All AEs and SAEs, will be followed until resolution, stabilization, the Investigator determines that it is no longer clinically significant, the event is otherwise explained, or the study participant is lost to follow-up (as defined in Section 7.3). Further information on follow-up procedures is given in Section 10.3.

8.3.4 Regulatory reporting requirements for SAEs

Prompt notification by the Investigator to the Sponsor of a SAE is essential so that legal obligations and ethical responsibilities towards the safety of participants and the safety of a study treatment under clinical investigation are met.

The Sponsor has a legal responsibility to notify both the local regulatory authority and other regulatory agencies about the safety of a study treatment under clinical investigation. The Sponsor will comply with country-specific regulatory requirements relating to safety reporting to the regulatory authority, IRB/IEC, and Investigators.

Investigator safety reports must be prepared for suspected unexpected serious adverse reactions (SUSAR) according to local regulatory requirements and Sponsor policy and forwarded to Investigators as necessary.

An Investigator who receives an Investigator safety report describing a SAE or other specific safety information (eg, summary or listing of SAEs) from the Sponsor will review and then file it along with the Investigator's Brochure and will notify the IRB/IEC, if appropriate according to local requirements.

8.3.5 Pregnancy

Details of all pregnancies in female participants and female partners of male participants who become pregnant will be collected after the start of study treatment and through the SFU visit (ie, 20 weeks after last dose of IMP).

If a pregnancy is reported, the Investigator must immediately inform the Sponsor within 24 hours of learning of the pregnancy and should follow the procedures outlined in Appendix 4.

A female study participant should be withdrawn from the study as soon as pregnancy is known (by positive pregnancy test), and the following should be completed:

- The study participant should return for an early discontinuation visit.
- The study participant should immediately stop the intake of the IMP.
- An SFU Visit should be scheduled 20 weeks after the study participant has received her last dose of IMP.

Abnormal pregnancy outcomes (eg, spontaneous abortion, fetal death, stillbirth, congenital anomalies, ectopic pregnancy) are considered SAEs.

8.3.6 Adverse events of special interest

An AE of special interest is any AE that a regulatory authority has mandated be reported on an expedited basis, regardless of the seriousness, expectedness, or relatedness of the AE to the administration of a UCB product/compound. Potential Hy's Law cases, defined as $\geq 3 \times \text{ULN}$ ALT or AST with coexisting $\geq 2 \times \text{ULN}$ total bilirubin in the absence of $\geq 2 \times \text{ULN}$ ALP, with no alternative explanation for the biochemical abnormality, must ALWAYS be reported to UCB as an AE of special interest (ie, without waiting for any additional etiologic investigations to have been concluded). Follow-up information should then be reported if an alternative etiology is identified during investigation and monitoring of the participant.

8.3.7 Other safety topics of interest

Prespecified safety topics of interest for the study are infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behavior, major adverse cardiovascular events, hepatic events and PDILI, malignancies, and inflammatory bowel disease.

These are based on findings from the IMP clinical program to date, potential risks generally associated with biologic immunomodulators, or findings from other medicines with a related mechanism of action. There are no specific AE reporting requirements for these topics, except those listed below for events relating to TB; however, special monitoring, additional data collection activities, and/or enhanced signal detection activities (within UCB) are in place.

The reporting requirements for events relating to TB are as follows:

- The IGRA test conversions defined as a positive or indeterminate (and confirmed indeterminate on repeat) should be reported as AEs. The AE term would need to be updated with final diagnosis once available.
- Latent TB infection must be reported as an AE. Follow-up reports should be completed as per protocol requirement until the LTBI resolves.
- Confirmed active TB is always considered an SAE and must be reported per SAE reporting instruction in the study protocol. Follow-up reports should be completed as per protocol requirement until TB infection resolves.

8.3.8 Anticipated serious adverse events

The following list of Anticipated SAEs (Table 8-1) is predicted to occur in the population studied in this protocol at some frequency that is independent of drug exposure. This original list will remain in effect for the duration of the protocol. Note that listed events will not be regarded as anticipated SAEs if they are life threatening or if they result in the death of the study participant.

This list does not change the Investigator's obligation to report all SAEs (including Anticipated SAEs) as detailed in Section 8.3.1 and Section 10.3.

Table 8–1: Anticipated SAEs for the Population of Participants with HS

MedDRA System Organ Class	MedDRA Preferred Term
Gastrointestinal disorders	Crohn's disease Colitis ulcerative
Psychiatric disorders	Depression Anxiety
Musculoskeletal and connective tissue disorders	Arthropathy
Skin and subcutaneous tissue disorders	Pyoderma gangrenosum Pilonidal cyst Acne conglobate Hidradenitis
Neoplasms benign, malignant and unspecified (including cysts and polyps)	Lymphoma Squamous cell carcinoma of skin
Infections and infestations	Cellulitis
Metabolism and nutritional disorders	Diabetes mellitus Dyslipidaemia Metabolic syndrome
Endocrine disorders	Thyroid disorder Polycystic ovaries

HS=hidradenitis suppurativa; MedDRA=Medical Dictionary for Regulatory Activities; SAE=serious adverse event

8.3.9 Suspected transmission of an infectious agent

For the purposes of reporting, any suspected transmission of an infectious agent via a medicinal product should be considered as an SAE; such cases must be reported immediately, recorded in the AE module of the eCRF, and followed as any other SAE. Any organism, virus, or infectious particle (eg, prion protein transmitting transmissible spongiform encephalopathy), pathogenic or nonpathogenic, is considered an infectious agent.

8.4 Safety signal detection

Selected data from this study will be reviewed periodically to detect as early as possible any safety concern(s) related to the IMP so that Investigators, clinical study participants, regulatory authorities, and IRBs/IECs will be informed appropriately and as early as possible.

The UCB study physician or medically qualified designee/equivalent will conduct an ongoing review of SAEs and perform ongoing SAE reconciliations in collaboration with the Patient Safety (PS) representative.

As appropriate for the stage of development and accumulated experience with the IMP, medically qualified personnel at UCB may identify additional safety measures (eg, AEs, vital signs, laboratory or ECG results) for which data will be periodically reviewed during the course of the study.

An independent DMC will be used in this study; see Section 9.7 and Section 10.1.5 for details.

8.5 Treatment of overdose

For this study, any dose of IMP greater than that prescribed in the protocol will be considered an overdose. Overdose events are only considered AEs or SAEs if there are associated clinical signs and symptoms or if the act of taking the excess IMP itself is an AE or SAE (eg, suicide attempt).

UCB does not recommend specific treatment for an overdose. Any signs or symptoms of adverse reactions should be treated symptomatically as per standard care by the Investigator.

Bimekizumab will not be self-administered by the study participant.

In the event of an overdose, the Investigator should:

1. Contact the Medical Monitor immediately.
2. Closely monitor the study participant for any AE/SAE and laboratory abnormalities until they have resolved, have a stable sequelae, the Investigator determines that it is no longer clinically significant, or the study participant is lost to follow up.
3. Document the quantity of the excess dose as well as the duration of the overdose in the CRF.

Decisions regarding dose interruptions will be made by the Investigator in consultation with the Medical Monitor based on the clinical evaluation of the study participant.

8.6 Pharmacokinetics

Blood samples will be collected prior to dosing for measurement of plasma concentrations of bimekizumab at all timepoints described in Table 1–1. A total of 9mL will be collected at timepoints for which PK, anti-drug antibodies (ADAb), and neutralizing antibodies are all measured and 3mL will be collected at timepoints for which only PK is measured. Instructions for the collection and handling of biological samples will be provided by the Sponsor. The actual date and time (24-hour clock time) of each sample will be recorded.

Samples will be used to evaluate the PK of bimekizumab. Samples collected for analyses of bimekizumab plasma concentration may also be used to evaluate safety or efficacy aspects related to concerns arising during or after the study.

Genetic analyses will not be performed on these samples. Study participant confidentiality will be maintained. At visits during which blood samples for the determination of plasma concentrations of bimekizumab will be taken, 1 sample of sufficient volume can be used.

Drug concentration information that may unblind the study will not be reported to any blinded study personnel as long as the study remains blinded.

8.7 Genetics

For individuals consenting to the pharmacogenetic substudy, blood samples will be drawn for exploratory genetic/epigenetic analyses at the timepoints specified in Table 1–1. Collection of these samples will enable evaluation of genetics/epigenetics biomarkers relative to disease biology and progression, drug treatment and inflammatory and immune response processes. The nature and format of these tentative analyses will be determined when the results of the main study are made available. A separate ICF will be required for those study participants who agree to participate in the pharmacogenetics substudy. The substudy will be conducted only where

ethically accepted and authorized by the regulatory agencies. Refusal to participate in the substudy will not affect a study participant's ability to participate in the main study.

The samples will be stored at -80°C at the central biorepository for up to 20 years.

In the event of deoxyribonucleic acid (DNA) extraction failure, a replacement genetic blood sample may be requested from the participant. Signed informed consent will be required to obtain a replacement sample unless it was included in the original consent.

See Appendix 5 (Section 10.5) for information regarding genetic research. Details on processes for collection and shipment and destruction of these samples can be found in laboratory manual for this study. Detailed information on sample analysis will be provided in a bioanalytical report.

8.8 Pharmacodynamics

See Section 8.9.

8.9 Biomarkers

Where local regulations permit, blood samples will be drawn for exploratory ribonucleic acid, proteins and metabolites biomarker analysis at the timepoints specified in Table 1–1. Where local regulations permit, urine samples will be drawn for exploratory proteins and metabolites biomarker analysis at the timepoints specified in Table 1–1. Collection of these samples will enable evaluation of biomarkers relative to disease biology and progression, drug treatment and inflammatory and immune response processes.

Instructions pertaining to sample collection, processing, storage, labeling, and shipping are provided in the laboratory manual for this study. The nature and format of these tentative analyses will be determined at a later stage. The samples will be stored at the secure long-term storage facility selected by UCB for up to 20 years.

These samples will only be used to further our understanding of HS and/or how biomarkers, including genetic variation, may affect response or be affected by treatment with bimekizumab in HS.

8.9.1 Immunogenicity assessments

Blood samples for the measurement of ADAb and neutralizing antibodies will be collected. A total of 9mL will be collected at timepoints for which PK, ADAb, and neutralizing antibodies are all measured. Immunogenicity data will not be sent to the Investigator to protect the blinded nature of the treatment assignments and response.

Antibodies to bimekizumab will be evaluated in plasma samples collected from all participants according to the Schedule of Activities. Additionally, plasma samples should also be collected at the final visit from participants who discontinued IMP or were withdrawn from the study. These samples will be tested by the Sponsor or Sponsor's designee.

Plasma samples will be screened for antibodies binding to bimekizumab and the titer of confirmed positive samples will be reported. Other analyses may be performed to verify the stability of antibodies to bimekizumab and/or further characterize the immunogenicity of bimekizumab.

The detection and characterization of antibodies to bimekizumab will be performed using validated assay methods by or under the supervision of the Sponsor. All samples collected for detection of antibodies to bimekizumab will also be evaluated for bimekizumab plasma concentration to enable interpretation of the antibody data. Antibodies may be further characterized and/or evaluated for their ability to neutralize the activity of bimekizumab. Samples may be stored for a maximum of 20 years (or according to local regulations) following the last study participant's last visit for the study at a facility selected by the Sponsor to enable further analysis of immune responses to bimekizumab.

8.10 Medical resource utilization and health economics

Health-related outcomes and medical resource utilization will be collected as part of standard eCRF pages during the study (eg, concurrent medical procedures, concomitant medications, hospitalizations, WPAI-SHP).

8.11 Photography

At certain sites, where feasible, representative photographs of the changes in skin will be captured. Photographs will be anonymized. This is optional for study participants and requires a separate informed consent. A decision not to consent does not exclude the study participant from the study.

9 STATISTICAL CONSIDERATIONS

A description of statistical methods follows and will be described in more detail in the Statistical Analysis Plan (SAP).

9.1 Definition of analysis sets

9.1.1 Enrolled Set

The Enrolled Set will consist of all study participants who have given informed consent.

9.1.2 Randomized Set

The Randomized Set (RS) will consist of all randomized study participants.

9.1.3 Safety Set

The Safety Set will consist of all study participants who received at least 1 full or partial dose of IMP and will be used for the demographic, safety, and immunogenicity analyses.

9.1.4 Full Analysis Set

The Full Analysis Set (FAS) will consist of all study participants who received at least 1 dose (full or partial) of IMP and had a valid Baseline measurement and a post-Baseline measurement for abscess, inflammatory nodules, and draining tunnel counts.

9.1.5 Per-Protocol Set

The Per-Protocol Set will consist of all study participants in the FAS who had no important protocol deviations affecting the primary efficacy variable. Important protocol deviations will be predefined and study participants with important protocol deviations will be evaluated during ongoing data cleaning meetings prior to unblinding of the data.

9.1.6 Pharmacokinetics Per-Protocol Set

The Pharmacokinetics Per-Protocol Set will consist of study participants who received at least 1 full dose of bimekizumab and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the PK. The Pharmacokinetics Per-Protocol Set is defined separately for each of the treatment periods (ie, separately for the Initial Treatment Period and the Maintenance Treatment Period).

9.1.7 COVID-19 Free Set

The COVID-19 Free Set will consist of all study participants randomized into the study and who have no COVID-19 impact through Week 16. This analysis set will be used for sensitivity analysis of the primary efficacy endpoint.

9.2 General statistical considerations

All analyses will be performed using SAS® version 9.4 or later (SAS Institute, Cary, NC, US).

Descriptive statistics will be used to provide an overview of the Baseline, efficacy, and safety results. For categorical parameters, the number and percentage of study participants in each category will be presented by treatment group. The denominator for the percentages will be based on the number of study participants appropriate for the purpose of analysis. Unless otherwise noted, all percentages will be expressed to 1 decimal place. For continuous parameters, descriptive statistics will include n, mean, standard deviation, median, minimum, and maximum. Two-sided 95% confidence intervals, geometric means, and coefficient of variation will be presented for selected variables as appropriate.

Baseline for each assessment is defined as either the value obtained at Baseline or the last available value obtained prior to treatment administration (details to be specified in the SAP).

Formal statistical testing will be conducted for this study for the primary and secondary efficacy variables. Other efficacy variables will be summarized descriptively by treatment arm. P-values and confidence intervals may be produced for other or exploratory variables but will be interpreted as non-inferential (ie, nominal). Additionally, other analyses will be conducted as deemed appropriate and described in the SAP.

The primary treatment comparison for all formal statistical analyses of efficacy will be between bimekizumab and placebo.

9.3 Planned efficacy/outcome analyses

9.3.1 Analysis of the primary efficacy endpoint

The primary objective of this randomized, double-blind, placebo-controlled, multicenter, pivotal study in study participants with moderate to severe HS is to compare the efficacy of bimekizumab 320mg Q2W and bimekizumab 320mg Q4W with placebo at Week 16. For the purposes of Week 16 analyses, the bimekizumab treatment arms of 320mg Q2W/Q2W and bimekizumab 320mg Q2W/Q4W treatment groups will be pooled.

The primary and secondary efficacy analyses will be performed based on the RS.

The primary endpoint is the HiSCR₅₀ response at Week 16. The primary efficacy analysis will evaluate the composite estimand in the RS. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and early receipt of

systemic antibiotic rescue medication, or discontinuation of IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for the primary efficacy analysis:

1. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
2. Study participant-level outcome=HiSCR₅₀ at Week 16.
3. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, and not discontinuing study treatment due to an AE or lack of efficacy through Week 16. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation.
4. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (ie, receipt of systemic antibiotic rescue medication, or discontinuation of IMP due to an AE or lack of efficacy) will be imputed as non-response.

The statistical hypothesis for the HiSCR₅₀ response at Week 16 is that the conditional odds ratio for the HiSCR₅₀ response in the bimekizumab treatment group relative to the placebo group is equal to 1.

A logistic regression model will be used to assess the effect of bimekizumab vs placebo on HiSCR₅₀ response. The model will include a fixed effect for treatment. The stratification variables of Hurley stage and prior antibiotic use will be added to the model unless inappropriate. The odds ratio versus placebo, p-value (from Wald test), and confidence interval will be calculated.

To assess the impact of the COVID-19 pandemic on the primary efficacy endpoint analysis, additional sensitivity analyses will be performed as specified in Section 9.3.4.

9.3.2 Analysis of the secondary efficacy endpoints

The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment:

1. Proportion of study participants who achieve HiSCR₇₅ at Week 16. A similar analysis to the primary responder analysis will be performed.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
2. Proportion of study participants who experience at least 1 flare by Week 16, with flare defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline. A similar analysis to the primary responder analysis will be performed.
 - a. bimekizumab 320mg Q2W vs placebo

-
- b. bimekizumab 320mg Q4W vs placebo
3. Absolute change from Baseline in DLQI Total Score at Week 16. Analysis will be based on an analysis of covariance (ANCOVA) with treatment and stratification variables as fixed effects and the Baseline values as covariate.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
 4. Absolute change from Baseline in Skin Pain Score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HSSDD. Analysis will be based on an ANCOVA with treatment and stratification variables as fixed effects and the Baseline values as covariate.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
 5. Pain response (defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change) at Week 16. A similar analysis to the primary responder analysis will be performed.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo

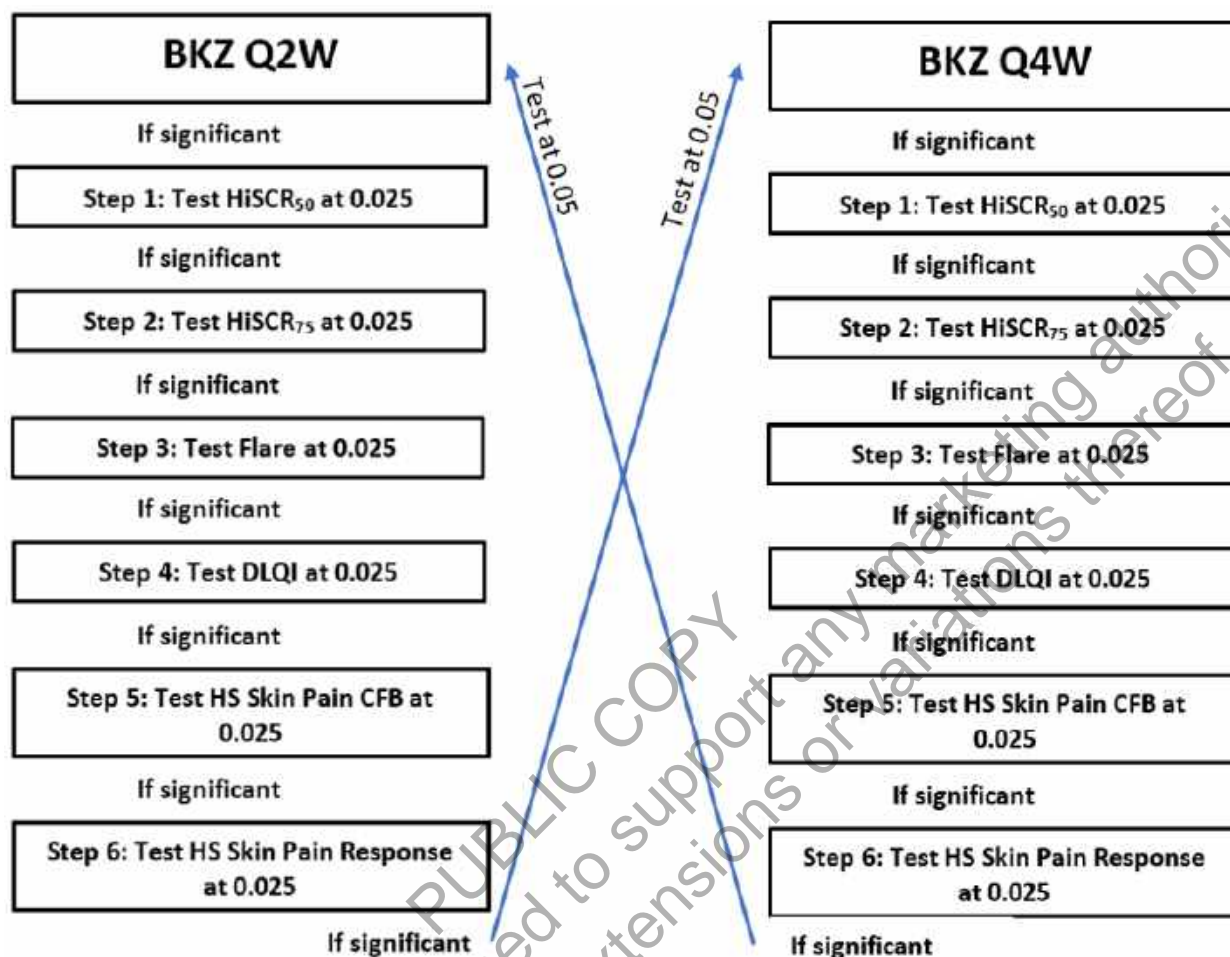
To control the overall type I error rate at 0.05 for the multiple comparisons in the primary and secondary efficacy variables, a closed testing procedure under a parallel gatekeeping framework will be applied ([Sun, 2018](#)).

Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$). Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

1. Step 1: Test $HiSCR_{50}$ at significance level 0.025.
2. Steps 2 to 6 – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown below, moving to the next step only if significance achieved at 0.025.
3. In the event that Step 6 is significant at 0.025 for a given dose, then Steps 1 to 6 will be repeated for the other dose using a significance level of 0.05.

A schematic of the procedure is shown in [Figure 9-1](#).

Figure 9-1: Closed Testing Procedure



AN=abscess and inflammatory nodule; CFB=change from Baseline; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks

9.3.3 Other efficacy/other outcome analyses

Analyses of the other efficacy measures will be detailed in the SAP.

9.3.4 COVID-19 impact analysis

To assess the impact of the COVID-19 pandemic on the primary efficacy endpoint analysis, the following additional sensitivity analyses will be performed:

- Imputation as nonresponse for any missing data due to COVID-19 at Week 16 and analyzed using the same analysis model as for the primary analysis
- Separate inferential analysis of the primary and secondary efficacy endpoints based on the COVID-19 Free Set

- Separate summary statistics for the primary efficacy endpoint at Week 16, based on the COVID-19 Free Set
- Summary of the number of study participants with missing primary or secondary endpoint data, as applicable

In addition, to assess the broader impact of the COVID-19 pandemic on the study, the following summaries will be presented:

- Summary of study participant disposition based on enrollment before, during, and after COVID-19 pandemic onset
- A summary of study visits impacted by the COVID-19 pandemic
- A summary of protocol deviations related to COVID-19

Any additional COVID-19 related analyses will be specified in the study SAP. Note: the date of COVID-19 pandemic onset is defined as 11 March 2020, the date that the World Health Organization declared the COVID-19 pandemic. The end date of the COVID-19 pandemic end will be defined in the study SAP, if applicable.

9.4 Planned safety and other analyses

9.4.1 Safety analyses

All TEAEs, SAEs, TEAEs leading to discontinuation, AEs of special interest (eg, cases meeting Hy's Law criteria), and other safety topics of interest (Section 8.3.7) will be collected during the study and for up to 20 weeks after the last dose of IMP (for study participants who do not participate in the extension study, HS0005). Safety analyses will be carried out using the Safety Set (study participants who received at least 1 full or partial dose of IMP). Summaries of Confirmed and Suspected COVID-19 TEAEs, respectively, will be presented. The definition of Confirmed and Suspected COVID-19 TEAEs will be provided in the SAP.

9.4.2 PK and ADAb analyses

Plasma concentrations of bimekizumab will be summarized by treatment group at each timepoint using descriptive statistics. In addition, PK model-based analyses may be performed.

Antidrug antibody data will be evaluated for each study participant and each regimen, and rates and classification of ADAb-positive study participants will be calculated.

9.5 Handling of protocol deviations

Important protocol deviations are identified as part of the data cleaning process as specified in the study-specific data cleaning schedule/plan. Ongoing, blinded data cleaning meetings will be held throughout the duration of the study. Objectives of these meetings include to review protocol deviations and to document potential impact that these deviations might have on the study objectives. Furthermore, overall trends in protocol deviations will be discussed at the Data Evaluation Meeting(s). Through this ongoing data cleaning and evaluation process, all decisions regarding important protocol deviations are made on an ongoing basis. Protocol deviations (eg, missing assessments or visits) related to COVID-19 will be listed and summarized separately.

9.6 Handling of dropouts or missing data

The analyses for the primary and secondary efficacy variables will include the use of multiple imputation. In multiple imputation, the missing value is replaced by a set of plausible values, where each value is a Bayesian draw from the conditional distribution of the missing data given the observed data.

Intermittent missing data will be imputed using the Markov-Chain Monte Carlo method, followed by regression for monotone missing data. The multiple imputation procedures planned for the primary and secondary efficacy analyses are based on an assumption of data missing at random.

The sensitivity of results to the approach for handling missing data will be evaluated via supportive analyses using different missing data mechanisms. The following supportive analyses for the primary efficacy variable will be conducted:

1. Deviations from the missing at random pattern will be evaluated using a reference-based multiple imputation approach. Intermittent missing data will be imputed using the Markov-Chain Monte Carlo method. The remaining monotone missing data will be assumed to follow a missing not at random pattern. These data will be imputed using a reference-based approach in which the multiple imputation model is based on data from the placebo group, thereby assuming that monotone missing data follow a trajectory similar to the placebo group.
2. Tipping point analyses will be performed to evaluate missingness assumptions. Various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. It will include scenarios where study participants who have missing data and are randomized to bimekizumab have a lower probability of response compared to study participants who have missing data and were randomized to placebo. For binary variables, this includes the worst case scenario where study participants who have missing data and are randomized to bimekizumab are considered nonresponders, while study participants who have missing data and were randomized to placebo are considered responders. The goal of the tipping point analysis is to systematically vary assumptions until there is no evidence of a treatment effect (if possible). The plausibility of such required delta adjustment will then be assessed.
3. The treatment policy strategy for addressing intercurrent events will be considered. This will be based on an analysis of all available data at Week 16 regardless of the occurrence of intercurrent events. This analysis will use the same models specified for the primary analysis, where study participants are analyzed according to their randomized treatment, even if they discontinued study treatment for any reason prior to Week 16. Even though efforts will be made to collect the primary outcome data for all study participants at Week 16, there may still be some study participants for whom Week 16 efficacy data cannot be obtained. In this case, missing data will be imputed using multiple imputation under the assumption of MAR. Results will be combined into a single inference using Rubin's rule. It should be noted that this measures something different from the primary analysis and could be confounded by placebo study participants who withdraw and are subsequently on another active medication at the time of the Week 16 assessment. Therefore, the results of this analysis should be interpreted in the appropriate context.

The analysis of the primary efficacy endpoint and secondary efficacy endpoints are based on a comparison of bimekizumab versus placebo at Week 16, with alpha adjustment strategy as indicated in Section 9.3.1 and Section 9.3.2.

The power to detect a statistically significant difference for each of the endpoints are shown in Table 9–1. Notably, with a 2-sided significance level of 0.025, the sample size of 140:70 provides 73% power for detecting at least a difference of 1.5 (bimekizumab Q4W vs placebo) for the Worst Pain change from Baseline (CFB) endpoint.

Given the high level of power for each of the primary and key secondary endpoints at the 0.025 significance level for the Q2W comparison ($\text{power} \geq 0.89$), and per the alpha spending strategy, there is a high likelihood that the Q4W comparison of Worst Pain CFB vs placebo will be allowed to be tested against the 0.05 level of significance. The power for this latter test is 81%. The sample size is thus ultimately driven by the Worst Pain CFB endpoint. Furthermore, the randomization ratio of 2:2:2:1 has been chosen to provide study participants with a high probability (6/7 ~ 86%) of being randomized to active study drug.

After randomization for this study was complete, an additional endpoint to assess Worst Pain response was included in the sequential testing procedure. This additional endpoint is defined as HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change at Week 16. Note that the power calculations reported in Table 9–1 for this endpoint are based on the sample size that was initially driven by the Worst Pain CFB endpoint as described above. With a 2-sided significance level of 0.025, the sample size of 104:52 in the subset of participants reporting Baseline HSSDD worst skin pain score at or beyond the threshold for clinically meaningful change (ie, Baseline HSSDD \geq threshold value) provides 53% power for detecting a statistically significant difference between bimekizumab Q4W and placebo in the proportion of Worst Pain responders.

Note that the power at the 0.025 level of significance associated with this endpoint for the comparison between bimekizumab Q2W and placebo is 95%. The Q4W comparison of Worst Pain response vs placebo against the 0.05 level of significance is therefore likely, and the power at this significance level is 65%. Given the strength of the power for the bimekizumab Q2W arm vs placebo across endpoints, it is considered acceptable to have a relatively low power for this final endpoint in the testing sequence for the bimekizumab Q4W treatment arm vs placebo

Table 9–1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
Flare	0.99	0.99	Proportion of participants with flare by Week 16=0.09	Proportion of participants with flare by Week 16=0.19	Proportion of participants with flare by Week 16=0.52
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6
Worst Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0004 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms.

^b Within-subject average of Worst Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Pain score at or above the threshold for clinically meaningful change from Baseline

10 SUPPORTING DOCUMENTATION AND OPERATIONAL CONSIDERATIONS

10.1 Appendix 1: Regulatory, ethical, and study oversight considerations

10.1.1 Regulatory and ethical considerations

The study will be conducted under the auspices of an IRB/IEC, as defined in local regulations, International Council for Harmonisation (ICH)- Good Clinical Practice (GCP), and in accordance with the ethical principles that have their origin in the Declaration of Helsinki.

The Investigator/UCB will ensure that an appropriately constituted IRB/IEC that complies with the requirements of the current ICH-GCP version or applicable country-specific regulations will be responsible for the initial and continuing review and approval of the clinical study. Prior to initiation of the study, the Investigator/UCB will forward copies of the protocol, ICF, Investigator's Brochure, Investigator's curriculum vitae (if applicable), advertisement (if applicable), and all other participant-related documents to be used for the study to the IRB/IEC for its review and approval.

Before initiating a study, the Investigator will have written and dated full approval from the responsible IRB/IEC for the protocol.

The Investigator will also promptly report to the IRB/IEC all changes in the study, all unanticipated problems involving risks to participants or others, and any protocol deviations, to eliminate immediate hazards to participants.

The Investigator will not make any changes in the study or study conduct without IRB/IEC approval, except where necessary to eliminate apparent immediate hazards to the participants. For minor changes to a previously approved protocol during the period covered by the original approval, it may be possible for the Investigator to obtain an expedited review by the IRB/IEC as allowed.

As part of the IRB/IEC requirements for continuing review of approved studies, the Investigator will be responsible for submitting periodic progress reports to the IRB/IEC (based on IRB/IEC requirements), at intervals appropriate to the degree of participant risk involved, but no less than once per year. The Investigator should provide a final report to the IRB/IEC following study completion.

UCB (or its representative) will communicate safety information to the appropriate regulatory authorities and all active Investigators in accordance with applicable regulatory requirements. The appropriate IRB/IEC will also be informed by the Investigator or the Sponsor, as specified by the applicable regulatory requirements in each concerned country. Where applicable, Investigators are to provide the Sponsor (or its representative) with evidence of such IRB/IEC notification.

10.1.2 Financial disclosure

Insurance coverage will be handled according to local requirements.

Finance and insurance are addressed in the Investigator and/or CRO agreements, as applicable.

10.1.3 Informed consent process

Study participant's informed consent must be obtained and documented in accordance with local regulations, ICH-GCP requirements, and the ethical principles that have their origin in the principles of the Declaration of Helsinki.

Prior to obtaining informed consent, information should be given in a language and at a level of complexity understandable to the study participant in both oral and written form by the Investigator (or designee). Each study participant will have the opportunity to discuss the study and its alternatives with the Investigator.

Prior to participation in the study, the ICF should be signed and personally dated by the participant, or his or her legal representative, and by the person who conducted the informed consent discussion (Investigator or designee). The study participant or his or her legal representative must receive a copy of the signed and dated ICF. As part of the consent process, each study participant must consent to direct access to his or her medical records for study-related monitoring, auditing, IRB/IEC review, and regulatory inspection.

If the ICF is amended during the study, the Investigator (or the Sponsor, if applicable) must follow all applicable regulatory requirements pertaining to the approval of the amended ICF by the IRB/IEC and use of the amended form.

All studies conducted at centers in the US must include the use of a Health Insurance Portability and Accountability Act Authorization form.

The study participant may withdraw his or her consent to participate in the study at any time. A study participant is considered as enrolled in the study when he or she has signed the ICF. A CRF must not be started, nor may any study specific procedure be performed for a given study participant, without having obtained his or her written consent to participate in the study.

10.1.4 Data protection

UCB staff (or designee) will affirm and uphold the study participant's confidentiality. Throughout this study, all data forwarded to UCB (or designee) will be identified only by the study participant number assigned at Screening.

The Investigator agrees that representatives of UCB, its designee, representatives of the relevant IRB/IEC, or representatives of regulatory authorities will be allowed to review that portion of the study participant's primary medical records that directly concerns this study (including, but not limited to, laboratory test result reports, ECG reports, admission/discharge summaries for hospital admissions occurring during a study participant's study participation, and autopsy reports for deaths occurring during the study).

The study participant must be informed that his or her personal study-related data will be used by the Sponsor in accordance with local data protection law. The level of disclosure must also be explained to the study participant.

The study participant must be informed that his or her medical records may be examined by Clinical Quality Assurance auditors or other authorized personnel appointed by the Sponsor, by appropriate IRB/IEC members, and by inspectors from regulatory authorities.

10.1.5 Committees structure

A DMC will be reviewing safety and efficacy data on an ongoing basis. The DMC membership includes experienced clinicians and a statistician, all of whom have expertise in clinical studies. Further details will be specified in the DMC Charter.

Cardiovascular, Gastrointestinal, and Neuropsychiatric Adjudication Committees will also periodically review data from this study. Details will be provided in the Cardiovascular, Gastrointestinal, and Neuropsychiatric Adjudication Committee charters.

Both DMC and Adjudication Committee members may not participate in the study as principal or co-Investigators, or as study participant care physicians and must not be members of the study team at UCB or the conducting CRO. The duration of membership for the committees will be inclusive of planned analyses for this study.

10.1.6 Data quality assurance

All study participant data relating to the study will be recorded on printed or electronic CRF unless transmitted to the Sponsor or designee electronically (eg, laboratory data). The Investigator is responsible for verifying that data entries are accurate and correct by physically or electronically signing the CRF.

The Investigator must maintain accurate documentation (source data) that supports the information entered in the CRF.

The Investigator must permit study-related monitoring, audits, IRB/IEC review, and regulatory agency inspections and provide direct access to source data documents.

The Sponsor or designee is responsible for the data management of this study including quality checking of the data.

Study monitors will perform ongoing source data verification to confirm that data entered into the CRF by authorized site personnel are accurate, legible, contemporaneous, original, and attributable from source documents; that the safety and rights of study participants are being protected; and that the study is being conducted in accordance with the currently approved protocol and any other study agreements, ICH-GCP, and all applicable regulatory requirements. Some study-specific assessments may be conducted remotely according to the study protocol during a pandemic or other exceptional circumstance (eg, hurricanes) due to the need to implement safety measures and guidance from regulatory authorities (refer to Section 8). Remote monitoring visits may be conducted during the COVID-19 pandemic or under other exceptional circumstances as deemed appropriate to ensure study participants' safety where local regulations permit.

All essential documents must be retained by the Investigator for the minimum retention period mandatory under the applicable local laws and regulations. The Investigator will contact UCB for authorization prior to the destruction of any study records or in the event of accidental loss or destruction of any study records. The Investigator will also notify UCB should he or she relocate or move the study-related files to a location other than that specified in the Sponsor's trial master file.

Quality tolerance limits will be established for the study using parameters related to patient safety reporting and reliability of study results. The parameters will be monitored throughout the

study to identify systematic issues. Parameters used, parameter values, important deviations from the quality tolerance limits, and actions taken will be summarized in the clinical study report.

10.1.6.1 Case report form completion

The Investigator is responsible for prompt reporting of accurate, complete, and legible data in the electronic CRFs and in all required reports.

Any change or correction to the CRF after saving must be accompanied by a reason for the change.

Corrections made after the Investigator's review and approval (by means of a password/electronic signature) will be reapproved by the Investigator.

The Investigator should maintain a list of personnel authorized to enter data into the electronic CRF.

Detailed instructions will be provided in the CRF Completion Guidelines.

10.1.7 Source documents

All source documents must be accurate, clear, unambiguous, permanent, and capable of being audited. They should be made using some permanent form of recording (ink, typing, printing, optical disc). They should not be obscured by correction fluid or have temporary attachments (such as removable self-stick notes). Photocopies and/or printouts of CRFs are not considered acceptable source documents.

Source documents are original records in which raw data are first recorded. These may include hospital/clinic/general practitioner records, charts, diaries, x-rays, laboratory results, printouts, pharmacy records, care records, ECG or other printouts, completed scales, quality of life questionnaires, or video, for example. Source documents should be kept in a secure, limited access area.

Electronic Patient-Reported Outcome (ePRO) measures (eg, DLQI, EQ-5D-3L, WPAI-SHP, TSQM-9, HiSQOL, Daily HS Symptom Diary, HS Symptom Questionnaire, PGI-S-HS, PGI-C-HS, PGI-S-SP, and PGI-C-SP) will be completed by each participant and will be collected electronically.

The data collection and database management system will be supplied by a vendor and will be compliant with the relevant regulations. The data collected on the ePROs will be uploaded to a central server database and will be sent electronically to UCB (or a designated CRO).

Source documents that are computer generated and stored electronically must be printed for review by the monitor (eg, ECG reports). Once printed, these copies should be signed and dated by the Investigator and become a permanent part of the study participant's source documents. The Investigator will facilitate the process for enabling the monitor to compare the content of the printout and the data stored in the computer to ensure all data are consistent.

Electronic data records must be saved and stored as instructed by UCB (or designee).

10.1.8 Study and Site Closure

The Sponsor designee reserves the right to close the study site or terminate the study at any time for any reason at the sole discretion of the Sponsor. Study sites will be closed upon study

completion. A study site is considered closed when all required documents and study supplies have been collected and a study-site closure visit has been performed.

The Investigator may initiate study-site closure at any time, provided there is reasonable cause and sufficient notice is given in advance of the intended termination.

Reasons for the early closure of a study site by the Sponsor or Investigator may include but are not limited to:

- Failure of the Investigator to comply with the protocol, the requirements of the IRB/IEC or local health authorities, the Sponsor's procedures, or GCP guidelines
- Inadequate recruitment of participants by the Investigator
- Discontinuation of further bimekizumab development

10.1.9 Publication Policy

The results of this study may be published or presented at scientific meetings. If this is foreseen, the Investigator agrees to submit all manuscripts or abstracts to the Sponsor before submission. This allows the Sponsor to protect proprietary information and to provide comments.

The Sponsor will comply with the requirements for publication of study results. In accordance with standard editorial and ethical practice, the Sponsor will generally support publication of multicenter studies only in their entirety and not as individual site data. In this case, a coordinating Investigator will be designated by mutual agreement.

Authorship will be determined by mutual agreement and in line with International Committee of Medical Journal Editors authorship requirements.

10.2 Appendix 2: Clinical laboratory tests

- The tests detailed in the table below will be performed by the central laboratory.
- Protocol-specific requirements for inclusion or exclusion of participants are detailed in Section 5.1 and Section 5.2 of the protocol.
- Additional tests may be performed at any time during the study as determined necessary by the Investigator or required by local regulations.

Protocol-Required Safety Laboratory Assessments

Laboratory Assessments	Parameters			
Hematology	Platelet Count	<u>RBC Indices:</u> Mean corpuscular hemoglobin Mean corpuscular hemoglobin concentration Mean corpuscular volume		<u>White Blood Cell Count with Differential:</u> Neutrophils Lymphocytes Atypical lymphocytes Monocytes Eosinophils Basophils
	RBC Count			
	Hemoglobin			
	Hematocrit			
Clinical Chemistry	Blood urea nitrogen	Potassium	Alanine aminotransferase ^a	Total and direct bilirubin ^{a,b}
	Creatinine	Sodium	Aspartate aminotransferase ^a	Glucose (record fasting or nonfasting in CRF)
	Bicarbonate	Calcium	Alkaline phosphatase	Gamma glutamyltransferase
	Uric acid	Chloride	Magnesium	hs-CRP ^c
	Lactate dehydrogenase	Total cholesterol		
Routine Urinalysis	Specific gravity pH, glucose, protein, ketones, nitrite, blood by dipstick Microscopic examination (if blood or protein is abnormal)			
Other Screening Tests	Pregnancy testing ^d Follicle stimulating hormone ^e Urine drug screen (amphetamines, barbiturates, cocaine, opiates, cannabinoids, and benzodiazepines) Serology (human immunodeficiency virus, Hepatitis B, Hepatitis C)			

Protocol-Required Safety Laboratory Assessments

ALT=alanine aminotransferase; AST=aspartate aminotransferase; CRF=case report form;

hs-CRP=high-sensitivity C-reactive protein; RBC=red blood cell; ULN=upper limit of normal

^a Details of liver chemistry stopping criteria and required actions and follow-up assessments after liver stopping or monitoring event are given in Section 7.1.2 and Appendix 6 (Section 10.6). All events of $\geq 3 \times$ ULN ALT or AST with coexisting $\geq 2 \times$ ULN total bilirubin in the absence of $\geq 2 \times$ ULN alkaline phosphatase, with no alternative explanation for the biochemical abnormality may indicate severe liver injury (possible Hy's Law) and must be reported as an SAE (excluding studies of hepatic impairment or cirrhosis).

^b If total bilirubin is $>$ ULN, a direct bilirubin estimation (%) will be performed.

^c hs-CRP will be tested at specified visits (Table 1-1).

^d A serum pregnancy test will be performed at Screening for all women of childbearing potential. A urine pregnancy test (urine dipstick analyzed locally) is also required at the Baseline Visit and all other visits in the Schedule of Activities (Table 1-1). Pregnancy test results must be negative prior to administering IMP.

^e A high follicle stimulating hormone level in the postmenopausal range may be used to confirm a postmenopausal state in women not using hormonal contraception or hormonal replacement therapy. However, in the absence of 12 months of amenorrhea, a single follicle stimulating hormone measurement is insufficient.

Investigators must document his or her review of each laboratory safety report.

Laboratory and/or analyte results (eg, hs-CRP, immunogenicity, PK) that could unblind the study will not be reported to any blinded study personnel as long as the study remains blinded.

10.3 Appendix 3: Adverse events – definitions and procedures for recording, evaluating, follow-up, and reporting

Definition of AE

AE Definition
<ul style="list-style-type: none">• An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of IMP, whether or not considered related to the IMP.• Note: An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of IMP.

Events Meeting the AE Definition
<ul style="list-style-type: none">• Any abnormal laboratory test results (hematology, clinical chemistry, or urinalysis) or other safety assessments (eg, ECG, radiological scans, vital signs measurements), including those that worsen from Baseline, considered clinically significant in the medical and scientific judgment of the Investigator (ie, not related to progression of underlying disease).• Exacerbation of a chronic or intermittent pre-existing condition including either an increase in frequency and/or intensity of the condition.• New conditions detected or diagnosed after IMP administration even though it may have been present before the start of the study.• Signs, symptoms, or the clinical sequelae of a suspected drug-drug interaction.• Signs, symptoms, or the clinical sequelae of a suspected overdose of either IMP or a concomitant medication. Overdose per se will not be reported as an AE/SAE unless it is an intentional overdose taken with possible suicidal/self-harming intent. Such overdoses should be reported regardless of sequelae.• "Lack of efficacy" or "failure of expected pharmacological action" per se will not be reported as an AE or SAE. Such instances will be captured in the efficacy assessments. However, the signs, symptoms, and/or clinical sequelae resulting from lack of efficacy will be reported as AE or SAE if they fulfil the definition of an AE or SAE.

Events NOT Meeting the AE Definition

- Any clinically significant abnormal laboratory findings or other abnormal safety assessments which are associated with the underlying disease, unless judged by the Investigator to be more severe than expected for the participant's condition.
- The disease/disorder being studied or expected progression, signs, or symptoms of the disease/disorder being studied, unless more severe than expected for the study participant's condition.
- Medical or surgical procedure (eg, endoscopy, appendectomy): the condition that leads to the procedure is the AE.
- Situations in which an untoward medical occurrence did not occur (social and/or convenience admission to a hospital).
- Anticipated day-to-day fluctuations of pre-existing disease(s) or condition(s) present or detected at the start of the study that do not worsen.

Definition of SAE

If an event is not an AE per definition above, then it cannot be an SAE even if serious conditions are met (eg, hospitalization for signs/symptoms of the disease under study, death due to progression of disease).

A SAE is defined as any untoward medical occurrence that, at any dose:
a. Results in death
b. Is life-threatening The term 'life-threatening' in the definition of 'serious' refers to an event in which the participant was at risk of death at the time of the event. It does not refer to an event, which hypothetically might have caused death, if it were more severe.
c. Requires inpatient hospitalization or prolongation of existing hospitalization In general, hospitalization signifies that the participant has been detained (usually involving at least an overnight stay) at the hospital or emergency ward for observation and/or treatment that would not have been appropriate in the physician's office or outpatient setting. Complications that occur during hospitalization are AEs. If a complication prolongs hospitalization or fulfills any other serious criteria, the event is serious. When in doubt as to whether "hospitalization" occurred or was necessary, the AE should be considered serious. Hospitalization for elective treatment of a pre-existing condition that did not worsen from Baseline is not considered an AE.
d. Results in persistent disability/incapacity <ul style="list-style-type: none">• The term disability means a substantial disruption of a person's ability to conduct normal life functions.• This definition is not intended to include experiences of relatively minor medical significance such as uncomplicated headache, nausea, vomiting, diarrhea, influenza, and accidental trauma (eg, sprained ankle) which may interfere with or prevent everyday life functions but do not constitute a substantial disruption.
e. Is a congenital anomaly/birth defect
f. Important medical events: <ul style="list-style-type: none">• Medical or scientific judgment should be exercised in deciding whether SAE reporting is appropriate in other situations such as important medical events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the participant or may require medical or surgical intervention to prevent one of the other outcomes listed in the above definition. These events should usually be considered serious.• Examples of such events include, but are not limited to, potential Hy's law, invasive or malignant cancers, intensive treatment in an emergency room or at home for allergic bronchospasm, blood dyscrasias or convulsions that do not result in hospitalization, or development of drug dependency or drug abuse.

Recording and Follow-Up of AE and/or SAE

AE and SAE Recording

- When an AE/SAE occurs, it is the responsibility of the Investigator to review all documentation (eg, hospital progress notes, laboratory reports, and diagnostics reports) related to the event.
- The Investigator will then record all relevant AE/SAE information in the CRF.
- It is **not** acceptable for the Investigator to send photocopies of the participant's medical records to UCB in lieu of completion of the AE/SAE CRF page.
- There may be instances when copies of medical records for certain cases are requested by UCB. In this case, all participant identifiers, with the exception of the participant number, will be redacted on the copies of the medical records before submission to UCB.
- The Investigator will attempt to establish a diagnosis of the event based on signs, symptoms, and/or other clinical information. Whenever possible, the diagnosis (not the individual signs/symptoms) will be documented as the AE/SAE.

Assessment of Intensity

The Investigator will make an assessment of intensity for each AE and SAE reported during the study and assign it to 1 of the following categories:

- Mild: An event that is easily tolerated by the participant, causing minimal discomfort and not interfering with everyday activities.
- Moderate: An event that causes sufficient discomfort and interferes with normal everyday activities.
- Severe: An event that prevents normal everyday activities. An AE that is assessed as severe should not be confused with a SAE. Severe is a category utilized for rating the intensity of an event; and both AEs and SAEs can be assessed as severe.
- An event is defined as 'serious' when it meets at least 1 of the predefined outcomes as described in the definition of an SAE, NOT when it is rated as severe (eg, a severe AE may be either serious or not serious, depending on whether these criteria are also met).

The National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE) should be used as a supportive standardization instrument to evaluate AEs and SAEs but the final intensity grading by the Investigator must be mild, moderate, or severe.

Assessment of Causality

- The Investigator is obligated to assess the relationship between IMP and each occurrence of each AE/SAE.
- A "reasonable possibility" of a relationship conveys that there are facts, evidence, and/or arguments to suggest a causal relationship, rather than a relationship cannot be ruled out.
- The Investigator will use clinical judgment to determine the relationship.
- Alternative causes, such as underlying disease(s), concomitant therapy, and other risk factors, as well as the temporal relationship of the event to IMP administration will be considered and investigated.
- The Investigator will also consult the Investigator's Brochure and/or Product Information, for marketed products, in his or her assessment.
- For each AE/SAE, the Investigator **must** document in the medical notes that he or she has reviewed the AE/SAE and has provided an assessment of causality.
- There may be situations in which an SAE has occurred and the Investigator has minimal information to include in the initial report to UCB. However, **it is very important that the Investigator always make an assessment of causality for every event before the initial transmission of the SAE data to UCB.**
- The Investigator may change his or her opinion of causality in light of follow-up information and send a SAE follow-up report with the updated causality assessment.
- The causality assessment is one of the criteria used when determining regulatory reporting requirements.

Follow-up of AEs and SAEs

- The Investigator is obligated to perform or arrange for the conduct of supplemental measurements and/or evaluations as medically indicated or as requested by UCB to elucidate the nature and/or causality of the AE or SAE as fully as possible. This may include additional laboratory tests or investigations, histopathological examinations, or consultation with other health care professionals.
- An AE should be followed until it has resolved, has a stable sequelae, the Investigator determines that it is no longer clinically significant, or the participant is lost to follow up. This follow-up requirement applies to AEs, SAEs, and AEs of special interest.
- If a participant dies during participation in the study or during a recognized follow-up period, the Investigator will provide UCB with a copy of any post-mortem findings including histopathology.
- New or updated information will be recorded in the originally completed CRF.
- The Investigator will submit any updated SAE data to UCB within 24 hours of receipt of the information.

Reporting of SAEs

SAE Reporting to UCB via an Electronic Data Collection Tool

- The primary mechanism for reporting an SAE to UCB will be the electronic data collection tool.
- If the electronic system is unavailable for more than 24 hours, then the site will use the paper SAE data collection tool (see next section).
- The site will enter the SAE data into the electronic system as soon as it becomes available.
- After the study is completed at a given site, the electronic data collection tool will be taken off-line to prevent the entry of new data or changes to existing data.
- If a site receives a report of a new SAE from a study participant or receives updated data on a previously reported SAE after the electronic data collection tool has been taken off-line, then the site can report this information on a paper SAE form (see next section) or to the UCB study physician by telephone.
- Contacts for SAE reporting can be found on the page after the title page of this protocol.

SAE Reporting to UCB via Paper CRF

- Facsimile transmission of the SAE paper CRF is the preferred method to transmit this information to the Medical Monitor or UCB study physician.
- In rare circumstances and in the absence of facsimile equipment, notification by telephone is acceptable with a copy of the SAE data collection tool sent by overnight mail or courier service.
- Initial notification via telephone does not replace the need for the Investigator to complete and sign the SAE CRF pages within the designated reporting time frames.
- Contacts for SAE reporting can be found on the page after the title page of this protocol.

10.4 Appendix 4: Contraceptive guidance and collection of pregnancy information

Definitions

Woman of Childbearing Potential (WOCBP)

A woman is considered fertile following menarche and until becoming post-menopausal unless permanently sterile (see below).

Women in the following categories **are not considered WOCBP**:

1. Premenarchal
2. Premenopausal female with 1 of the following:
 - Documented hysterectomy
 - Documented bilateral salpingectomy
 - Documented bilateral oophorectomy

Note: Documentation can come from the site personnel's review of the participant's medical records, medical examination, or medical history interview.

3. Postmenopausal female
 - A postmenopausal state is defined as no menses for 12 months without an alternative medical cause. A high follicle stimulating hormone level in the postmenopausal range may be used to confirm a postmenopausal state in women not using hormonal contraception or hormonal replacement therapy (HRT). However, in the absence of 12 months of amenorrhea, a single follicle stimulating hormone measurement is insufficient.
 - Females on HRT and whose menopausal status is in doubt will be required to use one of the non-estrogen hormonal highly effective contraception methods if they wish to continue their HRT during the study. Otherwise, they must discontinue HRT to allow confirmation of postmenopausal status before study enrollment.

Contraception Guidance

Female participants of childbearing potential are eligible to participate if they agree to use a highly effective method of contraception consistently and correctly as described in the table below.

Highly Effective Contraceptive Methods^a

Highly Effective Contraceptive Methods That Are User Dependent^b

Failure rate of <1% per year when used consistently and correctly.

Combined (estrogen and progestogen containing) hormonal contraception associated with inhibition of ovulation

- Oral
- Intravaginal
- Transdermal

Progestogen only hormonal contraception associated with inhibition of ovulation

- Oral
- Injectable

Highly Effective Methods That Are User Independent

Implantable progestogen only hormonal contraception associated with inhibition of ovulation

- Intrauterine device
- Intrauterine hormone-releasing system
- Bilateral tubal occlusion

Vasectomized partner

A vasectomized partner is a highly effective contraception method provided that the partner is the sole male sexual partner of the WOCBP and the absence of sperm has been confirmed. If not, an additional highly effective method of contraception should be used.

Sexual abstinence

Sexual abstinence is considered a highly effective method only if defined as refraining from heterosexual intercourse during the entire period of risk associated with the IMP. The reliability of sexual abstinence needs to be evaluated in relation to the duration of the study and the preferred and usual lifestyle of the participant.

NOTES:

- a) In case of newly started contraception pills/intrauterine devices, Investigator should consider the correct timing of starting/applying such methods in relation to the menstrual cycle and the manufacturing instruction as when these newly started methods would become effective.
- b) Typical use failure rates may differ from those when used consistently and correctly. Use should be consistent with local regulations regarding the use of contraceptive methods for participants participating in clinical studies.

Pregnancy testing

- WOCBP should only be included after a confirmed menstrual period and a negative highly sensitive serum pregnancy test.
- Additional pregnancy testing should be performed as indicated in the Schedule of Activities (Table 1–1) during the treatment period and at 20 weeks after the last dose of IMP and as required locally.
- Pregnancy testing will be performed whenever a menstrual cycle is missed or when pregnancy is otherwise suspected

Male participants with partners who become pregnant

- The Investigator will attempt to collect pregnancy information on any male participant's female partner who becomes pregnant while the male participant is in this study. This applies only to male participants who receive IMP. If the study participant is later found to be on placebo, then pregnancy data collection can stop.
- In cases where the partner of a male subject enrolled in a clinical study becomes pregnant, the Investigator or designee is asked to contact the subject to request consent of the partner via the Partner Pregnancy Consent form that has been approved by the responsible IRB/IEC and should be available in the Investigator site file. In case of questions about the consent process, the Investigator may contact the UCB/CRO contract monitor for the study. The Investigator will complete the information in the eCRF only after the partner has agreed that additional information can be captured and has provided the signed Partner Pregnancy Consent form. UCB's Patient Safety department is also the primary contact for any questions related to the data collection for the partner pregnancy, eventual birth, and follow up.
- After obtaining the necessary signed informed consent from the pregnant female partner directly, the Investigator will record pregnancy information on the appropriate form and submit it to the Sponsor within 1 working day of learning of the partner's pregnancy. The female partner will also be followed to determine the outcome of the pregnancy. Information on the status of the mother and child will be forwarded to the Sponsor. Generally, the follow-up will be at least 12 months after the delivery date. Any termination of the pregnancy will be reported regardless of fetal status (presence or absence of anomalies) or indication for the procedure.

Female Participants who become pregnant

- Any female participant who becomes pregnant while participating in the study will discontinue IMP or be withdrawn from the study.
- The Investigator will collect pregnancy information on any female participant who becomes pregnant while participating in this study. Information will be recorded on the appropriate form and submitted to the Sponsor within 1 working day of learning of a participant's pregnancy. The participant will be followed to determine the outcome of the pregnancy. The Investigator will collect follow-up information on the participant and the neonate and the information will be forwarded to the Sponsor. Generally, the follow-up will be at least 12 months after the delivery date. Any termination of pregnancy will be reported, regardless of fetal status (presence or absence of anomalies) or indication for the procedure.
- While pregnancy itself is not considered to be an AE or SAE, any pregnancy complication or elective termination of a pregnancy will be reported as an AE or SAE. A spontaneous abortion is always considered to be an SAE and will be reported as such. Any post-study pregnancy related SAE considered reasonably related to the IMP by the Investigator will be reported to the Sponsor as described in Section 8.3.5. While the Investigator is not obligated to actively seek this information in former study participants, he or she may learn of an SAE through spontaneous reporting.

10.5 Appendix 5: Genetics

Use and Analysis of DNA

Samples for potential future exploratory biomarker research will be collected and stored from consenting participants in the study. This sampling is optional for study participants and requires a separate informed consent. A decision not to consent does not exclude the study participant from the study:

- Blood sample for DNA.

These samples will only be used to further understanding of HS and/or how biomarkers, including genetic variation, may affect response or be affected by treatment with bimekizumab, background products, and/or concomitant medications in study participants with HS.

10.6 Appendix 6: Liver safety – suggested actions and follow-up assessments

Participants with potential drug-induced liver injury must be assessed to determine if IMP must be discontinued, as outlined in Section 7.1.2.

All PDILI events must be reported as an AE, and PDILI events meeting SAE criteria should be reported to the Sponsor within 24 hours of learning of the occurrence. Any PDILI event that meets the criterion for potential Hy's Law must be reported within 24 hours of learning of the occurrence as an AE of special interest (see Section 8.3.6), and, if applicable, also reported as an SAE (Section 8.3).

Evaluation of PDILI consists of the diagnostic testing and continued monitoring included in Table 10–1 (specific tests dependent on laboratory results and corresponding symptoms) and consultation with a local hepatologist (if applicable; discussed in Section 10.6.1). The local hepatologist is the expert usually consulted by the treating physician for assessment and management of potential hepatic disease. This would usually be a hepatologist, but may be a gastroenterologist. Additional investigation and monitoring may be required and adapted based on the diagnosis after the cause of the liver injury/abnormality is confirmed (details in Section 7.1.2.1.1).

Table 10–1: Required investigations and follow-up for PDILI

Laboratory value			Immediate		Follow up	
ALT or AST	Total bilirubin	Symptoms ^a of hepatitis or hypersensitivity	Consultation requirements	Actions	Testing	Evaluation
≥3×ULN	≥2×ULN ^b	NA	Hepatology consult ^c Medical Monitor must be notified within 24 hours (eg, by laboratory alert) and subject discussed with Medical Monitor ASAP.	Immediate IMP discontinuation. ^d	Essential: Must have repeat liver chemistry values and additional testing completed ASAP (see Section 10.6.3); recommended to occur at the site with HCP.	Monitoring of liver chemistry values at least twice per week until values normalize, stabilize, or return to within baseline values. ^e
≥3×ULN	NA	Yes				
≥8×ULN	NA	NA				

Table 10–1: Required investigations and follow-up for PDILI

Laboratory value			Immediate		Follow up	
ALT or AST	Total bilirubin	Symptoms ^a of hepatitis or hypersensitivity	Consultation requirements	Actions	Testing	Evaluation
≥5×ULN (and ≥2× Baseline) and <8×ULN	<2×ULN	No	Discussion with Medical Monitor required. Consider need for hepatology consult if there is no evidence of resolution (see Follow-up requirements). ^c	<p>Further investigation – immediate IMP discontinuation not required (see Section 10.6.2).</p> <p>IMP discontinuation required if any of the following occur:</p> <ul style="list-style-type: none"> • Subject cannot comply with monitoring schedule. • Liver chemistry values continue to increase • Liver chemistry values remain ≥5×ULN (and ≥2× baseline) after 4 weeks of monitoring without evidence of resolution 	<p>Essential: Every attempt must be made to have repeat liver chemistry values and additional testing completed within 48 hours at the site with HCP (see Section 10.6.3).</p>	<p>Monitoring of liver chemistry values at least twice per week for 2 weeks.^e</p> <ul style="list-style-type: none"> • Immediate IMP discontinuation required if liver chemistry values continue to increase. <p>After 2 weeks of monitoring liver chemistry values:</p> <ul style="list-style-type: none"> • ALT or AST remains ≥5×ULN <8×ULN, IMP should be temporarily withheld and subject should undergo repeat test in 2 weeks. <p>Continue IMP if ALT or AST values <5×ULN; continue to monitor at least twice per week until values normalize, stabilize, or return to within baseline values.</p> <p>If ALT or AST remains ≥5×ULN after second retest, immediate IMP discontinuation required.</p> <p>Continue to monitor until values normalize, stabilize, or return to within baseline values.^d</p>

Table 10–1: Required investigations and follow-up for PDILI

Laboratory value			Immediate		Follow up	
ALT or AST	Total bilirubin	Symptoms ^a of hepatitis or hypersensitivity	Consultation requirements	Actions	Testing	Evaluation

ALP=alkaline phosphatase; ALT=alanine aminotransferase; ASAP=as soon as possible; AST=aspartate aminotransferase; HCP=healthcare practitioner; IMP=investigational medicinal product; NA=not applicable; PDILI=potential drug induced liver injury; ULN=upper limit of normal

^a Hepatitis symptoms include fatigue, nausea, vomiting, and right upper quadrant pain or tenderness; hypersensitivity symptoms include eosinophilia (>5%), rash, and fever (without clear alternative cause).

^b If the subject also has $\geq 2 \times$ ULN ALP, the possibility of an indication of biliary obstruction should be discussed with the Medical Monitor.

^c Details provided in Section 10.6.1 . The local hepatologist is the expert usually consulted by the treating physician for assessment and management of potential hepatic disease. This would usually be a hepatologist, but may be a gastroenterologist.

^d Details are provided in Section 10.6.2.

^e Unless an alternative monitoring schedule is agreed by the Investigator and UCB responsible physician. Determination of stabilization is at the discretion of the Investigator in consultation with the hepatologist (as applicable) and UCB responsible physician, as needed.

The results of all monitoring, including laboratory testing and other testing, should be made available to the study site and Sponsor.

All initial tests resulting in abnormal hepatic laboratory values need to be repeated, but appropriate medical action must not be delayed waiting for the repeat result.

If tests are done locally for more rapid results, a concurrent sample should also be sent to the central laboratory whenever possible. Medical care decisions are to be made initially using the most rapidly available results and a conservative approach must be taken if the results from the 2 laboratory tests are significantly different. Data from the local and central laboratory are to be recorded on the applicable eCRF pages.

When IMP is discontinued, all concomitant medications and herbal supplements that are not medically necessary should also be discontinued. In these cases, the Investigator should also consider dose reduction for medically necessary concomitant medication and consider changing any medically required concomitant medication known to be hepatotoxic to a suitable alternative.

When IMP is stopped due to PDILI (as described in Section 7.1.2.1), IMP must be permanently discontinued unless a subsequent alternative diagnosis fully explains the hepatic findings. If a subsequent alternative diagnosis fully explains the hepatic findings, and the requirements provided in Section 10.6.2 are met, rechallenge with IMP may be appropriate.

The approach to investigate PDILI is summarized in Table 10-1.

10.6.1 Consultation with Medical Monitor and local hepatologist

Potential drug-induced liver injury events require notification of the Medical Monitor or UCB study physician within 24 hours (eg, by laboratory alert), and the study participant must be discussed with the Medical Monitor or UCB study physician as soon as possible. If required, the study participant must also be discussed with the local hepatologist. The local hepatologist is the expert usually consulted by the treating physician for assessment and management of potential hepatic disease. This would usually be a hepatologist, but may be a gastroenterologist. If determined necessary, this discussion should be followed by a full hepatology assessment (see Section 10.6.3) and SAE report (if applicable).

10.6.2 Immediate action: determination of IMP discontinuation

All PDILI events require immediate action, testing, and monitoring.

The immediate action is dependent on the laboratory values and symptoms of hepatitis or hypersensitivity and ranges from continuation of IMP (followed by immediate investigation) to immediate and permanent discontinuation (see Section 7.1.2 and Table 10-1 for details).

When IMP is discontinued, all concomitant medications and herbal supplements that are not medically necessary should also be discontinued. The Investigator should also consider dose reduction for medically necessary concomitant medication and consider changing any medically required concomitant medication known to be hepatotoxic to a suitable alternative.

10.6.2.1 IMP restart/rechallenge

Study participants who are immediately discontinued from IMP due to having met certain criteria for PDILI (as described in Section 7.1.2 and Table 10-1), but for whom an alternative

diagnosis is confirmed, ie, drug-induced liver injury is excluded, can rarely restart IMP. Rechallenge with IMP can occur only if ALL of the following requirements are met:

- The results of additional testing and monitoring described in Section 10.6.3 and Section 7.1.2.1.1 confirm a nondrug-related cause for the abnormal hepatic laboratory parameters and any associated symptoms (ie, a subsequent alternative diagnosis fully explains the hepatic findings).
- No alternative treatment options are available to the study participant.
- The study participant has shown clear therapeutic benefit from the IMP.
- Study participant's ALT or AST elevations do not exceed $\geq 3 \times \text{ULN}$.
- Study participant's total bilirubin is $< 1.5 \times \text{ULN}$.
- Study participant has no signs or symptoms of hypersensitivity.
- The rechallenge is approved by the UCB study physician and a hepatologist. The hepatologist must be external to UCB. It is recommended that the hepatologist be a local hepatology expert or the hepatologist treating the study participant.
- Study participant agrees to the Investigator-recommended monitoring plan and understands his or her individual benefit risk for restarting IMP and this is adequately documented.

10.6.3 Testing: identification/exclusion of alternative etiology

The measurements and additional information required for the assessment of PDILI events when there is a reasonable possibility that they may have been caused by the IMP are detailed in Table 10-2 (laboratory measurements) and Table 10-3 (additional information). Results of the laboratory measurements and information collected are to be submitted to the Sponsor on the corresponding eCRF. If the medical history of the study participant indicates a requirement for other assessments not included below, these additional assessments should be completed and submitted, as applicable.

All blood samples should be stored, if possible. If tests are done locally for more rapid results, a concurrent sample must also be sent to the central laboratory.

Table 10–2: PDILI laboratory measurements

Virology-related	Hepatitis A IgM antibody
	HBsAg
	Hepatitis E IgM antibody
	HBcAb-IgM
	Hepatitis C RNA
	Cytomegalovirus IgM antibody
	Epstein-Barr viral capsid antigen IgM antibody (if unavailable, obtain heterophile antibody or monospot testing)
Immunology	Anti-nuclear antibody (qualitative and quantitative)
	Anti-smooth muscle antibody (qualitative and quantitative)
	Type 1 anti-liver kidney microsomal antibodies (qualitative and quantitative)
Hematology	Eosinophil count
Urinalysis	Urine drug screen ^a
Chemistry	Amylase
	Sodium, potassium, chloride, glucose, blood urea nitrogen, creatinine
	Total bilirubin, ALP, AST, ALT, gamma-glutamyltransferase, total cholesterol, albumin
	If total bilirubin $\geq 1.5 \times \text{ULN}$, obtain fractionated bilirubin to obtain % direct bilirubin
	Serum creatine phosphokinase and lactate dehydrogenase to evaluate possible muscle injury causing transaminase elevation
Additional	Prothrombin time/INR ^b
	Serum pregnancy test ^c
	PK sample

ALP=alkaline phosphatase; ALT=alanine aminotransferase; AST=aspartate aminotransferase; HBcAb-IgM=hepatitis B core antibody-IgM; HBsAg=hepatitis B surface antigen; IgM=immunoglobulin M; INR=international normalized ratio; PDILI=potential drug-induced liver injury; PK=pharmacokinetic; RNA=ribonucleic acid; ULN=upper limit of normal

^a Tests in addition to the specified analytes may be performed based on the Investigator's medical judgment and study participant history.

^b Measured only for study participants with ALT $> 8 \times \text{ULN}$, elevations in total bilirubin, and symptoms of hepatitis or hypersensitivity. Hepatitis symptoms include fatigue, nausea, vomiting, and right upper quadrant pain or tenderness; hypersensitivity symptoms include eosinophilia ($> 5\%$), rash, and fever (without clear alternative cause).

^c For women of childbearing potential.

Additional information to be collected is presented in [Table 10–3](#).

Table 10–3: PDILI information to be collected

New or updated information
<ul style="list-style-type: none"> Concomitant prescription and over-the-counter medications (eg, acetaminophen, herbal remedies, vitamins); dosages and dates should be included.
<ul style="list-style-type: none"> Pertinent medical history, including the following: <ul style="list-style-type: none"> History of liver disease (eg, autoimmune hepatitis, nonalcoholic steatohepatitis or other “fatty liver disease”) Adverse reactions to drugs Allergies Relevant family history or inheritable disorders (eg, Gilbert’s syndrome, alpha-1 antitrypsin deficiency) Recent travel Progression of malignancy involving the liver (Note: Metastatic disease to the liver, by itself, should not be used as an explanation for significant AST and/or ALT elevations.)
<ul style="list-style-type: none"> The appearance or worsening of clinical symptoms of hepatitis or hypersensitivity (eg, fatigue, nausea, vomiting, right upper quadrant pain or tenderness, decreased appetite, abdominal pain, jaundice, fever, or rash)
<ul style="list-style-type: none"> Recent clinically significant hypotension or hypoxemia with compromised cardiopulmonary function
<ul style="list-style-type: none"> Alcohol and illicit drug use
<ul style="list-style-type: none"> Results of liver imaging or liver biopsy, if done
<ul style="list-style-type: none"> Results of any specialist or hepatology consult, if done
<ul style="list-style-type: none"> Any postmortem/pathology reports

ALT=alanine aminotransferase; AST=aspartate aminotransferase; PDILI=potential drug-induced liver injury

**10.7 Appendix 7: Medical device AEs, Adverse device effects, SAEs,
and device deficiencies: definition and procedures for
recording, evaluating, follow-up, and reporting**

Not applicable to this study.

PUBLIC COPY

This document cannot be used to support any marketing authorization
application and any extensions or variations thereof.

10.8 Appendix 8: Rapid alert procedures

Not applicable to this study.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.9 Appendix 9: Country-specific requirements

Country-specific requirements will be provided separately, as applicable.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

10.10 Appendix 10: Abbreviations and trademarks

ADAb	anti-drug antibody
AE	adverse event
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AN	abscess and inflammatory nodule
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
CAT	computed axial tomography
COVID-19	coronavirus disease 2019
CPK	creatine phosphokinase
CRF	case report form
CRO	contract research organization
CTCAE	Common Terminology Criteria for Adverse Events
CXR	chest x-ray
DLQI	Dermatology Life Quality Index
DMC	Data Monitoring Committee
DNA	deoxyribonucleic acid
ECG	electrocardiogram
eCRF	electronic Case Report form
eC-SSRS	electronic Columbia-Suicide Severity Rating Scale
EQ-5D-3L	European Quality-of-Life 5 dimensions-3 level questionnaire
EudraCT	European Union Drug Regulating Authorities Clinical Trials
FAS	Full Analysis Set
GCP	Good Clinical Practice
HCV	hepatitis C virus
HiSCR	Hidradenitis Suppurativa Clinical Response
HiSCR ₂₅	a 25% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSCR ₅₀	a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSCR ₇₅	a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count

HiSCR ₉₀	a 90% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSCR ₁₀₀	a 100% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count
HiSQOL	Hidradenitis Suppurativa Quality of Life
HRT	hormonal replacement therapy
HS	hidradenitis suppurativa
hs-CRP	high-sensitivity C-reactive protein
HSSDD	Hidradenitis Suppurativa Symptom Daily Diary
HSSQ	Hidradenitis Suppurativa Symptom Questionnaire
IBD	inflammatory bowel disease
ICF	informed consent form
ICH	International Council for Harmonisation
IEC	Independent Ethics Committee
Ig	immunoglobulin
IGRA	interferon gamma release assay
IHS4	International Hidradenitis Suppurativa Severity score system
IL	interleukin
IMP	investigational medicinal product
IND	Investigational New Drug
IRB	Institutional Review Board
IRT	interactive response technology
LTBI	latent tuberculosis infection
<i>M.</i>	<i>Mycobacterium</i>
mAb	monoclonal antibody
NTM	nontuberculous mycobacterial
NRS	numeric rating scale
NSAID	nonsteroidal anti-inflammatory drug
PDILI	potential drug-induced liver injury
PEOT	Premature End of Treatment Visit
PGI-C-HS	Patient Global Impression of Change in Hidradenitis Suppurativa Severity
PGI-C-SP	Patient Global Impression of Change in Severity of Skin Pain

PGI-S-HS	Patient Global Impression of Hidradenitis Suppurativa Severity
PGI-S-SP	Patient Global Impression of Severity of Skin Pain
PHQ-9	Patient Health Questionnaire Depression Module-9
PK	pharmacokinetic(s)
PRN	as needed
PRO	patient-reported outcome
PSO	psoriasis
Q2W	every 2 weeks
Q4W	every 4 weeks
QOL	quality of life
RS	Randomized Set
SAE	serious adverse event
SAP	statistical analysis plan
sc	subcutaneous
SFU	Safety Follow-up
TEAE	treatment-emergent adverse event
TB	tuberculosis
TSQM-9	Treatment Satisfaction Questionnaire – Medication 9
ULN	upper limit of normal
WOCBP	woman of childbearing potential
WPAI-SHP	Workers Productivity and Activity Impairment Questionnaire: Specific Health Problem

10.11 Appendix 11: Protocol amendment history

The Protocol Amendment Summary of Changes Table for the current amendment (Amendment 4) is located directly before the Table of Contents.

Amendment 3 (09 Feb 2021)

Overall Rationale for the Amendment

The main reason for this protocol amendment is due to Regulatory Agency feedback and to provide procedural clarifications.

Section # and Name	Description of Change	Brief Rationale
Title page	<ul style="list-style-type: none"> National Clinical Trial (NCT) number has been added 	NCT number was received on 24 Jan 2020.
Serious adverse event reporting	<ul style="list-style-type: none"> Fax and email for Japan have been removed 	To align with safety reporting guidelines
1.1 Synopsis 2.1 Study rationale 2.2 Background	<ul style="list-style-type: none"> Updated description of bimekizumab and other minor edits Order of secondary efficacy endpoints aligned with closed testing procedure (Figure 9-1) 	Rationale is the same as the description
1.1 Synopsis (Treatment Groups and Duration)	<ul style="list-style-type: none"> Updated Baseline antibiotic therapy strata to remove the 30% cap on enrollment [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] [REDACTED] 	<p>Removed the Baseline antibiotic strata cap per US FDA advice (Baseline antibiotic use should coincide with the prevalence of antibiotic use for the disease)</p> <p>To align with the final Data Monitoring Committee Charter (DMC) Charter and DMC Statistical Analysis Plan (SAP)</p>
1.3 Schedule of activities	<ul style="list-style-type: none"> Extended the line for concomitant medications and adverse events to show they are collected through the Safety Follow-Up Visit Added an additional footnote that past medical history includes tobacco and alcohol use Added Note in footnotes that study assessments could be completed remotely in exceptional circumstances Footnote “n” (formally footnote “m”) has been updated regarding collection of Hurley Stage 	<p>To align collection of adverse events and concomitant medications with the protocol body text</p> <p>To include/clarify that tobacco and alcohol use is part of the medical history</p> <p>Provided operational flexibility to allow assessments to be collected remotely due to COVID-19 or other exceptional circumstance (eg, hurricanes)</p> <p>Clarification to make footnote “n” consistent with the table</p>

Section # and Name	Description of Change	Brief Rationale
		regarding collection of Hurley Stage
3 Objectives and endpoints Other	<ul style="list-style-type: none"> Order of secondary efficacy endpoints aligned with closed testing procedure (Figure 9-1) Other safety topics of interest endpoints were updated 	<p>With progressive development of bimekizumab and based on the ongoing review of emerging safety data, depression has been removed as a safety topic of interest. Nevertheless, depression will continue to be monitored as a safety parameter by the Patient Health Questionnaire 9 (PHQ-9) and will be captured via routine adverse event (AE) reporting during the study. This update is considered a procedural change.</p> <p>Rational for reordering the secondary efficacy endpoints is the same as the description.</p>
4.1 Overall design 4.2 Scientific rationale for study design	<ul style="list-style-type: none"> Updated Baseline antibiotic strata to remove the 30% cap on enrollment 	Removed the Baseline antibiotic strata cap per US FDA advice (Baseline antibiotic use should coincide with the prevalence of antibiotic use for the disease)
5.1 Inclusion criteria	<ul style="list-style-type: none"> Criteria #2 added text that diagnosis must be verifiable through medical notes and documentation Criteria #5 was edited for clarity, and added text that diagnosis/inadequate response must be verifiable through medical notes and documentation 	Clarification of criteria
5.2 Exclusion criteria	<ul style="list-style-type: none"> Criterion #16 text was updated to clarify use of the Screening Version of the electronic Columbia-Suicidality Severity Rating Scale (eC-SSRS) Criterion #19 was updated to exclude participants with prior use of an IL-17 biologic response modifier or participation in an IL-17 biologic response modifier study unless an appropriate washout period (within 6 months 	<p>Clarification of criterion #16 as it is actually collected and assessed</p> <p>Criterion #19 was modified to allow enrollment of a moderate to severe HS population with real-world prior use of other medications with appropriate washout periods</p>

Section # and Name	Description of Change	Brief Rationale
	prior to the Baseline Visit or 5 half-lives, whichever is greater) has been performed	
5.4 Screen failures	<ul style="list-style-type: none"> Added bullet explaining that study participants who require incision and drainage procedures for HS lesions are to be screen failed 	Updated the screen failure criteria
6.5.1 Permitted concomitant treatments	<ul style="list-style-type: none"> Wound care updated to add that use of wound care dressings will be recorded Added Lesion care Updated Baseline antibiotic strata to remove the 30% cap on enrollment 	Clarifications of and additions to allowed concomitant treatments Removed the Baseline antibiotic strata cap per US FDA advice (Baseline antibiotic use should coincide with the prevalence of antibiotic use for the disease)
6.5.2 Prohibited concomitant treatments	<ul style="list-style-type: none"> Added washout periods for systemic antibiotics if applicable; other biologics; IL-17, IL-12, and IL-23 inhibitors; and janus kinase inhibitors Filgotinib and Upadacitinib added under janus kinase inhibitors Text added to clarify topical drugs Added herbal medications for HS and a washout period Updated vaccine criteria Text added to clarify that medications listed are currently available medications, but the protocol will account for medicine approvals in a given class during the course of the study 	Clarification of prohibited medications/therapies and the criteria for their exclusion and to allow enrollment of a moderate to severe HS population with real-world prior use of other medications (current and future) with appropriate washout periods
7.1 Discontinuation of study medication	<ul style="list-style-type: none"> Added a paragraph to clarify procedures if a study participant tested positive for COVID-19 or a suspected COVID-19 infection 	Updated for the COVID-19 pandemic
7.1.2.1 PDILI discontinuation criteria	<ul style="list-style-type: none"> Removed 'and permanent' from the first sentence and a crossreference to Section 10.6.2.1 has been included 	To be consistent with the PDILI criteria throughout the protocol and Appendix 10.6
7.1.3 Treatment interruptions	<ul style="list-style-type: none"> Text was updated to clarify that doses of IMP that were missed due to a reasonable interfering AE that does not allow administration of IMP due to safety reasons, will not be considered for the evaluation of study participant discontinuation. 	<p>Clarification of study procedures.</p> <p>In line with the exclusion criterion regarding infections, a specific infection-related IMP</p>

Section # and Name	Description of Change	Brief Rationale
	<p>It is still used to calculate compliance.</p> <ul style="list-style-type: none"> A specific infection-related IMP interruption criterion has been added 	interruption criterion was added to clarify that participants with serious or recurrent infections not responding to standard therapies are not exposed to immunomodulatory therapies until their infection has resolved. This is in line with most biologic therapies, including other anti-IL17s.
7.2 Participant discontinuation/withdrawal from the study	<ul style="list-style-type: none"> Added cross-reference to pregnancy section 	Clarification of study procedures
8 Study Assessments and Procedures	<ul style="list-style-type: none"> Added text about allowing study assessments to be performed remotely during a pandemic or other exceptional circumstance 	Updated to allow study to proceed during COVID-19, and other exceptional circumstances (eg, hurricanes)
8.1.1 Lesion count	<ul style="list-style-type: none"> Removed “hypertrophic” from description of scars and added “HS lesions” 	Clarification of lesion definition
8.1.4.6 Euro-Quality of Life 5-Dimensions, 3 levels	<ul style="list-style-type: none"> Removed sentences #2 and #3 regarding summary index scores Added a clarification to the last sentence 	Sentence removed in line with SAP update as index scores are not required for the clinical study report
8.2.2 Vital signs	<ul style="list-style-type: none"> Noncontact forehead added to body temperature measurement 	Updated for the COVID-19 pandemic and to align with the electronic Case Report Form (eCRF)
8.2.6.1 Assessments at Screening	<ul style="list-style-type: none"> Specified that the TB questionnaire is administered by the Investigator or their designee 	Clarification of study procedure
8.2.6.3.3 Tuberculosis questionnaire	<ul style="list-style-type: none"> Specified that the questionnaire is administered by the Investigator or their designee 	Clarification of study procedure
8.3 Adverse events and serious adverse events	<ul style="list-style-type: none"> Added statement that cases of COVID-19 infection will be recorded as AEs (or SAEs , as required) 	Updated for the COVID-19 pandemic
8.3.5 Pregnancy	<ul style="list-style-type: none"> Bulletpoint #2 “or be down-titrated as instructed at the early discontinuation visit” was deleted 	Down titration of bimekizumab is not required if IMP needs to be discontinued
8.3.7 Other safety topics of interest	<ul style="list-style-type: none"> Deleted depression from list of AEs considered safety topics of interest Updated “liver function test changes/enzyme elevations” to 	With progressive development of bimekizumab and based on the ongoing review of emerging safety data, depression has been

Section # and Name	Description of Change	Brief Rationale
	<p>“hepatic events and potential drug-induced liver injury (PDILI)”</p> <ul style="list-style-type: none"> Clarified major “adverse” cardiovascular events 	<p>removed as a safety topic of interest. Nevertheless, depression will continue to be monitored as a safety parameter by the PHQ-9 and will be captured via routine AE reporting during the study. This update is considered a procedural change</p> <p>To align with UCB internal documents regarding assessment of hepatic events and PDILI</p> <p>Typographical error corrected for major adverse cardiovascular events</p>
9.1.7 COVID-19 Free Set	<ul style="list-style-type: none"> Addition of subsection for a COVID-19 Free analysis set 	Updated for the COVID-19 pandemic
9.3.1 Analysis of the primary efficacy endpoint	<ul style="list-style-type: none"> Cross reference added to Section 9.3.4 for additional sensitivity analyses related to the assessment of COVID-19 pandemic on the primary efficacy endpoint analysis 	Updated for the COVID-19 pandemic
9.3.2 Analysis of secondary efficacy endpoints	<ul style="list-style-type: none"> Points #3 and #4 have been reordered to be consistent with Figure 9-1 (Closed Testing Procedure). Missing figure caption was added 	Rational is the same as the description
9.3.4 COVID-19 impact analysis	<ul style="list-style-type: none"> Addition of subsection for COVID-19 impact analysis 	Updated for the COVID-19 pandemic
9.4.1 Safety analysis	<ul style="list-style-type: none"> Addition of sentence that Summaries of Confirmed and Suspected COVID-19 TEAEs, respectively, will be presented and their definitions will be provided in the SAP 	Updated for the COVID-19 pandemic
9.5 Handling of protocol deviations	<ul style="list-style-type: none"> Added statement that COVID-related protocol deviations would be listed and summarized separately 	Updated for the COVID-19 pandemic
9.7.1 Data Monitoring Committee	<ul style="list-style-type: none"> [REDACTED] 	To align with the final DMC Charter and DMC SAP

Section # and Name	Description of Change	Brief Rationale
10.1.6 Data quality assurance	<ul style="list-style-type: none"> Added statement about performing some study-specific assessments remotely under certain exceptional circumstances 	Updated for the COVID-19 pandemic and other exceptional circumstances (eg, hurricanes)
10.4 Appendix 4: Contraceptive guidance and collection of pregnancy information	<ul style="list-style-type: none"> Deleted footnote “c” regarding hormonal contraception. “Hormonal contraception may be susceptible to interaction with the IMP, which may reduce the efficacy of the contraceptive method. In this case, 2 highly effective methods of contraception should be utilized during the treatment period and for at least 20 weeks after the last dose of IMP” 	To be consistent with other studies in the bimekizumab program
10.6 Appendix 6: Liver safety – suggested actions and follow-up assessments	<ul style="list-style-type: none"> Paragraph #2 has been updated to add that PDILI events meeting SAE criteria should be reported to the Sponsor within 24 hours of learning of the occurrence. The requirement to report to the study site has been removed. 	Correction of typographical error and clarification of reporting procedures
10.11 Appendix 11: Protocol amendment history	<ul style="list-style-type: none"> Updated with changes from previous global amendment 	Self-evident
Throughout	<ul style="list-style-type: none"> Minor editorial and formatting revisions 	Minor edits and formatting revisions that do not impact content were made for readability and/or clarity

Amendment 2 (16 Dec 2019)

Overall Rationale for the Amendment

The main reason for global protocol amendment 2 was to update the study discontinuation/withdrawal criteria for study participants with IBD.

Section # and Name	Description of Change	Brief Rationale
7.2 Participant discontinuation/withdrawal from the study	<ul style="list-style-type: none"> Added text to IBD discontinuation/withdrawal criteria 	Previously approved text relating to discontinuation/withdrawal criteria for IBD was inadvertently removed from the original version of the protocol dated

Section # and Name	Description of Change	Brief Rationale
		29 Oct 2019. It is now being replaced
7.2 Participant discontinuation/withdrawal from the study	<ul style="list-style-type: none"> Fixed numbering in list of criteria 	Corrected typographical error in list numbering
10.11 Appendix 11: Protocol amendment history	<ul style="list-style-type: none"> Updated with changes from previous global protocol amendment 1 	Updated

Amendment 1 (06 Dec 2019)

Overall Rationale for the Amendment

The main reason for global protocol amendment 1 was to update the company name in line with the new Code of Companies and Associations recently adopted by Belgium.

Section # and Name	Description of Change	Brief Rationale
Title page	<ul style="list-style-type: none"> Updated company name from UCB Biopharma SPRL to UCB Biopharma SRL 	Change in company name on 02 Dec 2019
3 Objectives and Endpoints	<ul style="list-style-type: none"> Clarified wording of exploratory biomarker objective 	Clarification of objective with current genetics and biomarkers sections (Section 8.7 and Section 8.9, respectively)
8.2.6.3.7 Tuberculosis management of LTBI, active TB, or other NTB infection identified during study	<ul style="list-style-type: none"> Figure 8-1 was updated to reflect Screening terminology as follows: <ul style="list-style-type: none"> Green ovals that said “subject eligible for study continuation” or “subject eligible to continue the study” in original protocol were changed to say “subject eligible for study” Red ovals that said “withdraw from the study” were changed to say “subject NOT eligible for study” 	Figure 8-1 was corrected to reflect Screening terminology
10.11 Appendix 11: Protocol amendment history	<ul style="list-style-type: none"> Stated location of summary of changes table for the current amendment 	Updated

11 REFERENCES

Alikhan A, Lynch PJ, Eisen DB. Hidradenitis suppurativa: a comprehensive review. *J Am Acad Dermatol*. 2009;60(4):539-61.

Basra MK, Chowdhury MM, Smith EV, Freemantle N, Piquet V. A review of the use of the dermatology life quality index as a criterion in clinical guidelines and health technology assessments in psoriasis and chronic hand eczema. *Dermatol Clin*. 2012;30(2):237-44.

Bharmal M, Payne K, Atkinson MJ, Desrosiers MP, Morisky DE, Gemmen E. Validation of an abbreviated Treatment Satisfaction Questionnaire for Medication (TSQM-9) among patients on antihypertensive medications. *Health Qual Life Outcomes*. 2009;7:36.

Cho KA, Suh JW, Lee KH, Kang JL, Woo SY. IL-17 and IL-22 enhance skin inflammation by stimulating the secretion of IL-1 β by keratinocytes via the ROS-NLRP3-caspase-1 pathway. *Int Immunol*. 2012;24(3):147-58.

Cosentyx [package insert]. East Hanover, NJ: Novartis Pharmaceuticals Corporation; 2019. <https://www.pharma.us.novartis.com/sites/www.pharma.us.novartis.com/files/cosentyx.pdf><https://www.pharma.us.novartis.com/sites/www.pharma.us.novartis.com/files/cosentyx.pdf>.

Danby FW, Margesson LJ. Hidradenitis suppurativa. *Dermatol Clin*. 2010;28(4):779-93.

Esmann S, Dufour DN, Jemec GB. Questionnaire-based diagnosis of hidradenitis suppurativa: specificity, sensitivity and positive predictive value of specific diagnostic questions. *Br J Dermatol*. 2010;163(1):102-6.

Finlay AY. Quality of life assessments in dermatology. *Semin Cutan Med Surg*. 1998;17(4):291-6.

Frew JW, Hawkes JE, Krueger JG. A systematic review and critical evaluation of immunohistochemical associations in hidradenitis suppurativa. *F1000Research*. 2018;7.

Humira [package insert]. North Chicago, IL: AbbVie Inc; 2019. <https://www.rxabbvie.com/pdf/humira.pdf>.

Hurley H. Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In: *Dermatologic surgery*. Roenigh R, Roenigh H, editors. Marcel Dekker. 1989;729-39.

Kimball AB, Kerdel F, Adams D, Mrowietz U, Gelfand JM, Gniadecki R. Adalimumab for the treatment of moderate to severe Hidradenitis suppurativa: a parallel randomized trial. *Ann Intern Med*. 2012;157(12):846-55.

Kimball AB, Jemec GB, Yang M, Kageliery A, Signorovitch JE, Okun MM, et al. Assessing the validity, responsiveness and meaningfulness of the Hidradenitis Suppurativa Clinical Response (HiSCR) as the clinical endpoint for hidradenitis suppurativa treatment. *Br J Dermatol*. 2014;171(6):1434-42.

Kimball AB, Okun MM, Williams DA, Gottlieb AB, Papp KA, Zouboulis CC, et al. Two Phase 3 Trials of Adalimumab for Hidradenitis Suppurativa. *N Engl J Med*. 2016a;375(5):422-34.

Kimball AB, Sobell JM, Zouboulis CC, Gu Y, Williams DA, Sundaram M, et al. HiSCR (Hidradenitis Suppurativa Clinical Response): a novel clinical endpoint to evaluate therapeutic

outcomes in patients with hidradenitis suppurativa from the placebo-controlled portion of a phase 2 adalimumab study. *J Eur Acad Dermatol Venereol*. 2016b;30(6):989-94.

McMillan, K. Hidradenitis suppurativa: number of diagnosed patients, demographic characteristics, and treatment patterns in the United States. *Am J Epidemiol*. 2014;179(12):1477-83.

Martorell A, García-Martínez FJ, Jiménez-Gallo D, Pascual JC, Pereyra-Rodríguez J, Salgado L, Vilarrasa E. An update on hidradenitis suppurativa (part I): epidemiology, clinical aspects, and definition of disease severity. *Actas Dermo-Sifiliográficas (English Edition)*. 2015 Nov 1;106(9):703-15.

Mundt JC, Greist JH, Gelenberg AJ, Katzelnick DJ, Jefferson JW, Modell JG. Feasibility and validation of a computer-automated Columbia-Suicide Severity Rating Scale using interactive voice response technology. *J Psychiatr Res*. 2010;44(16):1224-8.

Naldi L. Epidemiology. In: Hidradenitis Suppurativa. Chapter 8. Jemec GB, Revuz J, Leyden J, editors. Springer-Verlag. 2006:58-64.

Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry*. 2011;168(12):1266-77.

Prens E, Deckers I. Pathophysiology of hidradenitis suppurativa: an update. *J Am Acad Dermatol*. 2015;73(5):S8-11.

Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993;4(5):353-65.

Revuz JE, Canoui-Poittrine F, Wolkenstein P, Vialette C, Gabison G, Pouget F, et al. Prevalence and factors associated with hidradenitis suppurativa: results from two case-control studies. *J Am Acad Dermatol*. 2008;59(4):596-601.

Riis PT, Sørensen K, Saunte DM, Jemec GB. Patients with hidradenitis suppurativa carry a higher systemic inflammatory load than other dermatological patients. *Archives of dermatological research*. 2015 Dec 1;307(10):885-9.

Riis PT, Vinding GR, Ring HC, Jemec GB. Disutility in patients with hidradenitis suppurativa: a cross-sectional study using EuroQoL-5D. *Acta Derm Venereol*. 2016;96(2):222-6.

Sartorius K, Emtestam L, Jemec GB, Lapins J. Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *Br J Dermatol*. 2009;161(4):831-9.

Schlapbach C, Hänni T, Yawalkar N, Hunger RE. Expression of the IL-23/TH17 pathway in lesions of hidradenitis suppurativa. *J Am Acad Dermatol*. 2011;65(4):790-8.

Sun H, Snyder E, Koch GG. Statistical planning in confirmatory clinical trials with multiple treatment groups, multiple visits, and multiple endpoints. *J Biopharmaceut Stat*. 2018;28(1):189-211.

Thorlacius L, Ingram JR, Villumsen B, et al. A core domain set for hidradenitis suppurativa trial outcomes: an international Delphi process. *Br J Dermatol*. 2018.

Van der Zee HH, de Ruiter L, Van Den Broecke DG, Dik WA, Laman JD, Prens EP. Elevated levels of tumour necrosis factor (TNF)- α , interleukin (IL)-1 β and IL-10 in hidradenitis suppurativa skin: a rationale for targeting TNF- α and IL-1 β . *British Journal of Dermatology*. 2011 Jun;164(6):1292-8.

Vender R, Lynde C, Ho V, Chau D, Poulin-Costello M. Work productivity and healthcare resource utilization outcomes for patients on etanercept for moderate to severe plaque psoriasis: Results from a 1-year, multicentre, open-label, single-arm study in a clinical setting. *Appl Health Econ Health Policy*. 2012;10(5):343-53.

von der Werth JM, Jemec GB. Morbidity in patients with hidradenitis suppurativa. *Br J Dermatol*. 2001;144:809-13.

Wolkenstein P, Loundou A, Barrau K, Auquier P, Revuz J; Quality of Life Group of the French Society of Dermatology. Quality of life impairment in hidradenitis suppurativa: a study of 61 cases. *J Am Acad Dermatol*. 2007;56:621-3.

Zouboulis CC, Desai N, Emtestam L, Hunger RE, Loannides D, Juhász I, et al. European S1 guideline for the treatment of hidradenitis suppurativa/acne inversa. *J Eur Acad Dermatol Venereol*. 2015;29(4):619-44.

Zouboulis CC, Tzellos T, Kyrgidis A. Development and validation of IHS4, a novel dynamic scoring system to assess hidradenitis suppurativa/acne inversa severity. *Br J Dermatol*. 2017;177:1401-9.

SPONSOR DECLARATION

I confirm that I have carefully read and understand this protocol and agree to conduct this clinical study as outlined in this protocol and according to current Good Clinical Practice.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Approval Signatures

Name: HS0004-protocol-amendment-4

Version: 1 . 0

Document Number: CLIN-000188473

Title: HS0004-Protocol Amendment 4

Approved Date: 08 May 2022

Document Approvals	
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 06-May-2022 16:01:58 GMT+0000
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 06-May-2022 16:48:41 GMT+0000
Approval Verdict: Approved	Name: [REDACTED] Capacity: Medical Date of Signature: 08-May-2022 20:16:36 GMT+0000

STATISTICAL ANALYSIS PLAN

Study: HS0003

Product: Bimekizumab

A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED,
MULTICENTER STUDY EVALUATING THE
EFFICACY AND SAFETY OF BIMEKIZUMAB IN STUDY
PARTICIPANTS WITH MODERATE TO SEVERE HIDRADENITIS SUPPURATIVA

SAP/Amendment Number	Date
Final SAP	3 Apr 2020
SAP Amendment 1	9 Dec 2021
SAP Amendment 2	11 Jul 2022
SAP Amendment 3	07 Oct 2022

Confidentiality Statement

Confidential

This document is the property of UCB and may not – in full or in part – be passed on, reproduced, published, or otherwise used without the express permission of UCB.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS	9
1 INTRODUCTION	13
2 PROTOCOL SUMMARY	13
2.1 Study objectives	13
2.1.1 Primary objective	13
2.1.2 Secondary objectives	13
2.1.3 Other objectives	13
2.1.4 Exploratory objective	13
2.2 Study endpoints	14
2.2.1 Efficacy endpoints	14
2.2.1.1 Primary efficacy endpoint	14
2.2.1.2 Secondary efficacy endpoints	14
2.2.1.3 Other efficacy endpoints	14
2.2.2 Pharmacokinetic and pharmacogenomic endpoints	16
2.2.2.1 Pharmacokinetic endpoints	16
2.2.2.2 Exploratory pharmacogenomic endpoints	16
2.2.3 Safety endpoints	16
2.2.3.1 Secondary safety endpoints	16
2.2.3.2 Other safety endpoints	16
2.2.4 Immunological endpoints	17
2.3 Study design and conduct	17
2.3.1 Study description	17
2.3.2 Study periods	18
2.3.2.1 Screening Period	18
2.3.2.2 Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)	18
2.3.2.3 Safety Follow-up Visit	18
2.3.3 Study duration per participant	18
2.3.4 Planned number of participants and sites	19
2.3.5 Anticipated regions and countries	19
2.4 Determination of sample size	19
3 DATA ANALYSIS CONSIDERATIONS	20
3.1 General presentation of summaries and analyses	20
3.2 General study level definitions	22
3.2.1 Relative day	22
3.2.2 Mapping of data from Premature End of Treatment visits	23
3.3 Definition of Baseline values	23

3.4	Protocol deviations.....	24
3.5	Analysis sets.....	24
3.5.1	Enrolled Set	24
3.5.2	Randomized Set	24
3.5.3	Safety Set	24
3.5.4	Full Analysis Set.....	24
3.5.5	Active Medication Set	24
3.5.6	Maintenance Set.....	24
3.5.7	Per-Protocol Set	24
3.5.8	Pharmacokinetics Per-Protocol Set.....	25
3.5.9	COVID-19 Free Set	25
3.6	Treatment assignment and treatment groups	25
3.7	Center pooling strategy	25
3.8	Coding dictionaries	26
3.9	Definition of an intercurrent event.....	26
3.10	Changes to protocol-defined analyses	27
4	STATISTICAL/ANALYTICAL ISSUES	27
4.1	Adjustments for covariates	27
4.2	Handling of dropouts or missing data.....	28
4.2.1	Efficacy data	28
4.2.1.1	Handling missing data for the primary efficacy endpoint	28
4.2.1.2	Handling missing data for the secondary efficacy endpoints.....	28
4.2.1.3	Handling missing data for the other efficacy endpoints.....	28
4.2.1.4	Missing Data Overview and Summary.....	29
4.2.2	Missing data algorithms for efficacy analyses.....	30
4.2.2.1	MI – MCMC / Monotone Regression	31
4.2.2.2	MI – MCMC / Reference-based imputation.....	36
4.2.2.3	Tipping Point Analysis	37
4.2.3	Rationale for estimand.....	39
4.2.3.1	Composite estimand	39
4.2.3.2	Hypothetical estimand.....	39
4.2.4	Dates and times.....	40
4.3	Interim analyses and data monitoring.....	41
4.3.1	Data monitoring committee	41
4.3.2	Interim analysis.....	42
4.4	Multicenter studies.....	42
4.5	Multiple comparisons/multiplicity.....	42
4.6	Use of an efficacy subset of participants	44

4.7	Active-control studies intended to show equivalence.....	44
4.8	Examination of subgroups	44
5	STUDY POPULATION CHARACTERISTICS.....	45
5.1	Study participant disposition.....	45
5.2	Impact of COVID-19	46
5.3	Protocol deviations.....	46
6	DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS	46
6.1	Demographics	46
6.2	Other Baseline characteristics.....	47
6.3	Medical history and concomitant diseases.....	48
6.4	Prior and concomitant medications.....	48
6.4.1	Assignment of medications to study period.....	49
6.4.2	Classification of participants as analgesic, antihistamine users	50
7	MEASUREMENTS OF TREATMENT COMPLIANCE.....	50
8	EFFICACY ANALYSES	51
8.1	Lesion count assessment	51
8.2	Primary efficacy endpoint.....	52
8.2.1	Derivation of HiSCR ₅₀ at Week 16	56
8.2.2	Primary analysis of the primary efficacy endpoint.....	57
8.2.3	Sensitivity analyses of the primary efficacy endpoint.....	58
8.2.3.1	Nonresponse imputation.....	58
8.2.3.2	MI-MCMC / Reference-based imputation	58
8.2.3.3	Tipping point analysis	58
8.2.3.4	Treatment policy.....	59
8.2.3.5	Analysis on observed cases	59
8.2.3.6	Analysis on FAS.....	59
8.2.3.7	Analysis on PPS	59
8.2.3.8	Analysis on CFS	59
8.2.3.9	Analysis including COVID-19 impact as intercurrent event	59
8.2.3.10	Cochran-Mantel-Haenszel test	60
8.2.3.11	Center-by-Treatment Interaction	60
8.3	Secondary efficacy endpoints	61
8.3.1	HiSCR ₇₅ at Week 16.....	65
8.3.2	Change from Baseline in DLQI Total Score at Week 16	65
8.3.2.1	Primary analysis of change from Baseline in DLQI Total score at Week 16	67
8.3.2.2	Sensitivity analysis of change from Baseline in DLQI Total score at Week 16	67

8.3.3	Change from Baseline in Skin Pain score at Week 16, as assessed by the “worst skin pain” item in the HSSDD	67
8.3.3.1	Primary analysis of change from Baseline in skin pain score at Week 16	68
8.3.3.2	Sensitivity analysis of change from Baseline in skin pain score at Week 16	68
8.3.4	HSSDD skin pain response at Week 16	68
8.3.4.1	Primary analysis of skin pain response at Week 16	68
8.3.4.2	Sensitivity analyses of Skin Pain Response at Week 16	69
8.4	Other efficacy endpoints	69
8.4.1	HiSCR endpoints	69
8.4.1.1	HiSCR ₂₅ , HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀	69
8.4.1.2	Time to response of HiSCR ₂₅ , HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀	Error! Bookmark not defined.
8.4.1.3	HiSCR ₂₅ , HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀ response at both Weeks 16 and 48	71
8.4.1.4	HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀ among Week 16 Responders	71
8.4.1.5	Time to loss of response of HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀ in Week 16 responders	71
8.4.1.6	Partial response	72
8.4.2	Lesion count	72
8.4.2.1	Change from Baseline in lesion count	72
8.4.2.2	AN count of 0, 1, or 2	73
8.4.2.3	AN ₂₅ , AN ₅₀ , AN ₇₅ , AN ₉₀ , AN ₁₀₀	73
8.4.3	Flare by Week 16	73
8.4.4	Flare relative to Baseline	74
8.4.5	Time to flare by Week 16	74
8.4.6	Time to flare by Week 48	74
8.4.7	International Hidradenitis Suppurativa Severity Score System (IHS4)	75
8.4.8	HS-Physician’s Global Assessment 6-point scale	75
8.4.9	High Sensitivity C-Reactive Protein (hs-CRP)	78
8.4.10	Initiation of systemic antibiotic rescue therapy	79
8.4.11	Time to initiation of systemic rescue therapy in the Initial Treatment Period	79
8.4.12	Time to an intercurrent event in the Initial Treatment Period	79
8.4.13	Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)	80
8.4.14	Hidradenitis Suppurativa Symptom Questionnaire (HSSQ)	81
8.4.15	DLQI	82
8.4.16	Hidradenitis Suppurativa Quality of Life (HiSQOL)	82
8.4.17	Patient Global Impression of HS Severity (PGI-S-HS)	84
8.4.18	Patient Global Impression of Change in HS Severity (PGI-C-HS)	84

8.4.19	Patient Global Impression of Severity of Skin Pain (PGI-S-SP).....	84
8.4.20	Patient Global Impression of Change in Severity of Skin Pain (PGI-C-SP)	84
8.4.21	Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)	85
8.4.22	Work Productivity and Activity Impairment Questionnaire–Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores.....	85
8.4.23	Treatment Satisfaction Questionnaire – Medication-9	86
8.4.24	Lesion intervention	87
8.5	Additional statistical analyses of other efficacy endpoints.....	87
9	PHARMACOKINETICS AND PHARMACODYNAMICS	88
9.1	Pharmacokinetics	88
9.2	Pharmacodynamics	89
9.3	Immunogenicity	89
9.3.1	Autoantibodies	89
9.3.2	Anti-bimekizumab antibodies.....	89
10	SAFETY ANALYSES.....	93
10.1	Extent of exposure	93
10.1.1	Exposure during the Initial Treatment Period.....	94
10.1.1.1	Study medication duration (days).....	94
10.1.1.2	Time at risk (days).....	94
10.1.2	Exposure during the Maintenance Treatment Period	95
10.1.2.1	Study medication duration (days).....	95
10.1.2.2	Time at risk (days).....	95
10.1.3	Exposure during the Initial and Maintenance Treatment Period.....	96
10.1.3.1	Study medication duration (days).....	96
10.1.3.2	Time at risk (days).....	96
10.2	Adverse events	97
10.2.1	Data considerations.....	97
10.2.1.1	COVID-19 related considerations	99
10.2.2	AE summaries.....	99
10.2.3	Other Safety topics of interest	101
10.2.3.1	Infections (serious, opportunistic, fungal and TB).....	101
10.2.3.2	Malignancies	104
10.2.3.3	Major adverse cardiac event.....	104
10.2.3.4	Neutropenia	106
10.2.3.5	Suicidal Ideation and Behavior	106
10.2.3.6	Inflammatory bowel disease.....	107
10.2.3.7	Hypersensitivity (including anaphylaxis).....	108
10.2.3.8	Hepatic events and PDILI	109

10.3	Clinical laboratory evaluations	109
10.4	Vital signs, physical findings, and other observations related to safety	112
10.4.1	Vital signs	112
10.4.2	Electrocardiograms	113
10.4.3	Other safety endpoints	113
10.4.3.1	Physical examination.....	114
10.4.3.2	Columbia-Suicide Severity Rating Scale (C-SSRS)	114
10.4.3.3	Assessment and management of TB and TB risk factors.....	114
10.4.3.4	Pregnancy testing	115
10.4.3.5	Patient Health Questionnaire (PHQ)-9 scores	115
11	REFERENCES	116
12	APPENDICES	118
12.1	Appendix A: MedDRA algorithmic approach to anaphylaxis.....	118
12.2	Appendix B: Definition of CTCAE grades.....	120
13	AMENDMENT TO THE STATISTICAL ANALYSIS PLAN.....	122
13.1	Amendment 1.....	122
13.2	Amendment 2.....	175
13.3	Amendment 3.....	175
	STATISTICAL ANALYSIS PLAN SIGNATURE PAGE.....	216

LIST OF TABLES

Table 2-1:	Power calculation assumptions and methods.....	20
Table 3-1:	Geographic regions and corresponding countries.....	25
Table 4-1:	Missing data handling approach by endpoint priority and type.....	30
Table 4-2:	Imputation allowable ranges by variable	32
Table 8-1:	Estimand Details and Attributes for Primary Endpoint	52
Table 8-2:	Estimand Details and Attributes for Secondary Analyses	61
Table 8-3:	Dermatology Life Quality Index.....	66
Table 8-4:	HS-Physician's Global Assessment 6-point scale.....	76
Table 8-5:	Hidradenitis Suppurative Quality of Life	83
Table 10-1:	Cardiovascular event classifications.....	105
Table 10-2:	Neuropsychiatric event classifications.....	107
Table 10-3:	IBD event classifications	108
Table 10-4:	Definitions of Markedly Abnormal Liver Function Values	110
Table 10-5:	Definitions of Markedly Abnormal Biochemistry Values.....	111

Table 10–6:	Definitions of Markedly Abnormal Hematology Values.....	111
Table 10–7:	Definitions of Markedly Abnormal Blood Pressure Values	113
Table 12–1:	Definition of CTCAE grades by biochemistry parameters.....	120
Table 12–2:	Definitions of CTCAE grades by hematology parameter.....	120

LIST OF FIGURES

Figure 2-1:	Study Schematic.....	17
Figure 4-1:	Sequence of testing	43

LIST OF ABBREVIATIONS

List of Abbreviations

%ΔAN	percentage change from Baseline in abscess and inflammatory nodule count
ADAb	anti-bimekizumab antibodies
AE(s)	adverse event(s)
ALP	alkaline phosphatase
ALQ	above the limit of quantification
ALT	alanine aminotransferase
AMS	Active Medication Set
AN	abscess and inflammatory nodule
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BKZ	bimekizumab
BLQ	below the limit of quantification
CFB	change from Baseline
CFS	COVID-19 Free Set
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
COVID-19	coronavirus disease 2019
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CV	coefficient of variation
CV-CAC	Cardiovascular Event Adjudication Committee
DLQI	Dermatology Life Quality Index
DMC	data monitoring committee
eCDF	empirical cumulative distribution function
ECG	electrocardiogram
eCRF	electronic case report form
eC-SSRS	electronic Columbia-Suicide Severity Rating Scale
EAER	exposure adjusted event rate
EAIR	exposure adjusted incident rate

List of Abbreviations

EQ-5D-3L	European Quality of Life-5 Dimensions-3 Level questionnaire
ES	Enrolled Set
FAS	Full Analysis Set
FDA	Food and Drug Administration
geoCV	geometric coefficient of variation
GGT	gamma-glutamyltransferase
HiSCR	Hidradenitis Suppurativa Clinical Response
HiSQOL	Hidradenitis Suppurativa Quality of Life
HLT	high level term
HS	hidradenitis suppurativa
hs-CRP	high sensitivity C-reactive protein
HSSDD	Hidradenitis Suppurativa Symptom Daily Diary
HSSQ	Hidradenitis Suppurativa Symptom Questionnaire
IBD	Inflammatory bowel disease
IBD-CAC	Inflammatory Bowel Disease Adjudication Committee
ICF	Informed Consent Form
ICH	International Council for Harmonisation
IDC	Infectious Disease Committee
IGRA	interferon gamma release assay
IHS4	International Hidradenitis Suppurativa Severity Scoring System
IMP	investigational medicinal product
LFT	liver function tests
LLOQ	lower limit of quantification
LOCF	last observation carried forward
LSM	least square mean
MACE	major cardiovascular events
MAR	missing at random
MCID	minimal clinically important difference
MCMC	Markov-Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MI-MCMC	multiple imputation Markov-Chain Monte Carlo

List of Abbreviations

MS	Maintenance Set
MSR	minimum significant ratio
n	number of study participants
NAb	neutralizing antibody
NI	Negative Immunodepletion
nR	New Ratio
NRI	nonresponder imputation
NRS	numeric rating scale
NS	Negative Screen
OC	observed case
PD	pharmacodynamic(s)
pDILI	potential drug induced liver injury
PEOT	premature end of treatment
PGI-C-HS	Patient Global Impression of Change in Hidradenitis Suppurativa Severity
PGI-C-SP	Patient Global Impression of Change in Severity of Skin Pain
PGI-S-HS	Patient Global Impression of Hidradenitis Suppurativa Severity
PGI-S-SP	Patient Global Impression of Severity of Skin Pain
PHQ-9	Patient Health Questionnaire 9
PI	Positive Immunodepletion
PK	pharmacokinetic(s)
PK-PPS	Pharmacokinetic Per-Protocol Set
PPS	Per-Protocol Set
PS	Positive Screen
PT	preferred term
Q2W	every 2 weeks
Q4W	every 4 weeks
QTcF	QT corrected for heart rate using Fridericia's formula
RS	Randomized Set
SAP	statistical analysis plan
SD	standard deviation
SE	standard error

List of Abbreviations

SFU	Safety Follow-up
SIB	suicidal ideation and behavior
SMQ	standardized MedDRA query
SOC	system organ class
SS	Safety Set
SSD	Safety Signal Detection
TEAE	treatment-emergent adverse event
TEMA	treatment-emergent markedly abnormal
TSQM-9	Treatment Satisfaction Questionnaire – Medication 9
ULN	upper limit of normal
VAS	visual analogue scale
WHODD	World Health Organization Drug Dictionary
WPAI-SHP	Work Productivity and Activity Impairment Questionnaire-Specific Health Problem

1 INTRODUCTION

This statistical analysis plan (SAP) defines the scope of statistical analyses and provides a detailed description of statistical methodology to support the final clinical study report (CSR).

The SAP is based on the Protocol Amendment 5, 27 September 2022. If a future protocol amendment necessitates a substantial change to the statistical analysis of the study data, this SAP may be amended accordingly. The content of this SAP is compatible with the International Council for Harmonisation (ICH)/Food and Drug Administration (FDA) E9 Guidance documents.

2 PROTOCOL SUMMARY

2.1 Study objectives

2.1.1 Primary objective

The primary objective of this study is to evaluate the efficacy of bimekizumab in study participants with moderate to severe hidradenitis suppurativa (HS).

2.1.2 Secondary objectives

The secondary objectives of this study are to:

- Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS
- Evaluate the safety of bimekizumab in study participants with moderate to severe HS

2.1.3 Other objectives

The other objectives of this study are to:

- Evaluate the efficacy of bimekizumab on Hidradenitis Suppurativa Clinical Response (HiSCR), other HS Scores, and other clinical measures of disease activity at various timepoints in study participants with moderate to severe HS
- Evaluate the efficacy of bimekizumab on abscesses, nodules, and draining tunnels at various timepoints in study participants with moderate to severe HS
- Evaluate the efficacy of bimekizumab on patient-reported outcome measures at various timepoints in study participants with moderate to severe HS
- Evaluate the effect of bimekizumab on other safety measures at various timepoints in study participants with moderate to severe HS
- Evaluate the pharmacokinetics (PK) of bimekizumab in study participants with moderate to severe HS
- Evaluate the immunogenicity of bimekizumab (antidrug antibodies) in study participants with moderate to severe HS

2.1.4 Exploratory objective

The exploratory objective of the study is to evaluate biomarkers in study participants with moderate to severe HS.

2.2 Study endpoints

2.2.1 Efficacy endpoints

2.2.1.1 Primary efficacy endpoint

The primary efficacy endpoint is the HiSCR₅₀ (defined as at least a 50% reduction from Baseline in the total abscess and inflammatory nodule [AN] count with no increase from Baseline in abscess or draining tunnel count) at Week 16.

2.2.1.2 Secondary efficacy endpoints

The secondary efficacy endpoints are defined as:

- HiSCR₇₅ response (defined as at least a 75% reduction from Baseline in the total AN count with no increase from Baseline in abscess or draining tunnel count) at Week 16
- Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16
- Absolute change from Baseline (CFB) in Skin Pain score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HS Symptom Daily Diary (HSSDD)
- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline

2.2.1.3 Other efficacy endpoints

The other efficacy endpoints are defined as:

- Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀
- HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀
- Absolute change from Baseline in International Hidradenitis Suppurativa Severity Score System (IHS4)
- Change from Baseline in the HS-Physician’s Global Assessment 6-point scale
- Absolute and percentage change from Baseline in high-sensitivity C-reactive protein (hs-CRP)
- Initiation of systemic antibiotic rescue therapy
- HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ at both Weeks 16 and 48
- Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders
- Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₅₀ during the Maintenance Treatment Period
- Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period

- Change and percentage change from Baseline in lesion counts (abscess count, inflammatory nodule count, AN count, and draining tunnel count)
- AN count of 0, 1, or 2
- AN₂₅, AN₅₀, AN₇₅, AN₉₀, AN₁₀₀ (defined as a 25%, 50%, 75%, 90%, 100% reduction in the total AN count relative to Baseline)
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48
- Time to flare from Weeks 0 to 16
- Time to flare from Week 16 to 48
- Absolute and percentage change (worst and average skin pain) from Baseline in HS Skin Pain score (11-point numeric rating scale)
- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline
- Absolute change from Baseline in DLQI Total Score
- DLQI Total Score of 0 or 1
- Minimum clinically important difference (MCID) (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4
- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) and Total score
- Patient Global Impression of HS Severity (PGI-S-HS)
- Patient Global Impression of Change of HS Severity (PGI-C-HS)
- Patient Global Impression of Severity of Skin Pain (PGI-S-SP)
- Patient Global Impression of Change of Skin Pain (PGI-C-SP)

- Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor
- Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor
- Responses to the European Quality of Life-5 Dimensions-3 Level questionnaire (EQ-5D-3L), absolute and changes from Baseline in EQ-5D-3L visual analog scale (VAS) scores
- Absolute change from Baseline in Work Productivity and Activity Impairment Questionnaire-Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores
- Domain Scores (effectiveness, convenience and global satisfaction) on the Treatment Satisfaction Questionnaire – Medication-9 (TSQM-9)

2.2.2 Pharmacokinetic and pharmacogenomic endpoints

2.2.2.1 Pharmacokinetic endpoints

The PK endpoint is the plasma bimekizumab concentrations over the study duration.

2.2.2.2 Exploratory pharmacogenomic endpoints

[REDACTED]

A specific SAP will be written to describe the analysis methods for those endpoints, as the results will not be summarized in the CSR. The nature and format of these analyses will be detailed in this SAP.

2.2.3 Safety endpoints

2.2.3.1 Secondary safety endpoints

The secondary safety endpoints are

- Treatment-emergent Adverse Events (TEAEs)
- Serious TEAEs
- TEAEs leading to withdrawal from study

2.2.3.2 Other safety endpoints

The other safety endpoints are

- Adverse events of special interest (Hy's Law)
- Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behavior, major adverse cardiovascular events, hepatic events and potential drug-induced liver injury (PDILI), malignancies, and inflammatory bowel disease.

- Absolute change from Baseline in the Patient Health Questionnaire (PHQ-9) score
- Absolute change from Baseline in vital signs
- Absolute change from Baseline in clinical laboratory values (chemistry and hematology)
- Electrocardiogram (ECG) results

2.2.4 Immunological endpoints

The immunological endpoints are

- Bimekizumab antidrug antibodies
- Bimekizumab neutralizing antibodies

The results of the bimekizumab neutralizing antibody analysis will not be summarized in the CSR for this study. All neutralizing antibody analyses will be detailed in the integrated immunogenicity SAP.

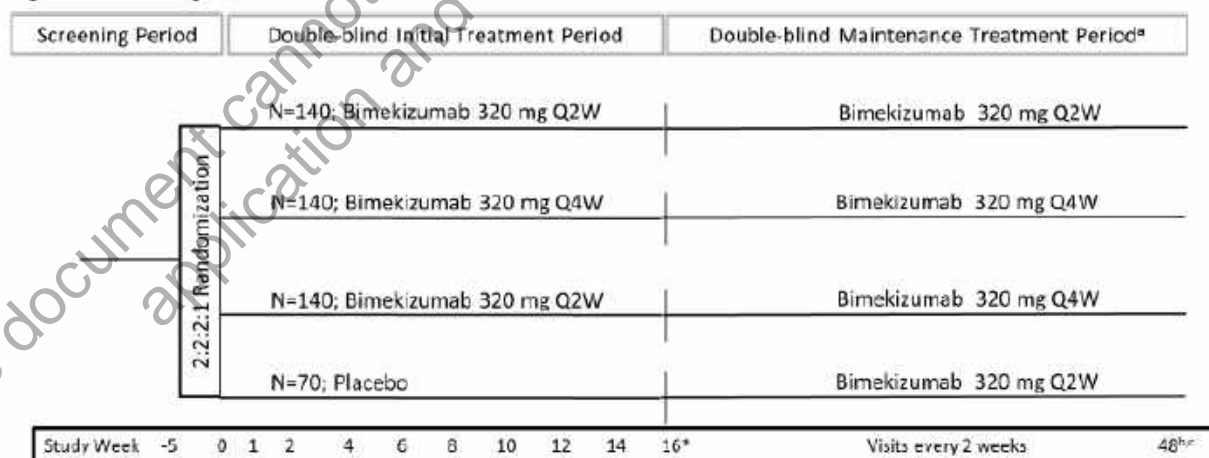
2.3 Study design and conduct

2.3.1 Study description

HS0003 is a Phase 3, randomized, double-blind, placebo-controlled, multicenter, pivotal study evaluating the efficacy and safety of bimekizumab in study participants with moderate to severe HS. Study participants meeting the inclusion criteria who do not meet any exclusion criteria will complete a Screening Period of 14 days to up to 5 weeks; a double-blind, 48-week Treatment Period comprising a 16-week Initial Treatment Period and 32-week Maintenance Treatment Period; and a 20-week Safety Follow-up (SFU) Period following the final injection of investigational medicinal product (IMP) if study participants do not enter a subsequent extension study (HS0005) or withdraw prematurely from treatment.

Study participants will be randomized in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to 1 of 3 dose regimens of bimekizumab or placebo as shown in the schematic (Figure 2-1).

Figure 2-1: Study Schematic



HiSCR₅₀=a 50% reduction in the total abscess and inflammatory nodule count with no increase from Baseline in abscess or draining tunnel count;

IMP=investigational medicinal product; Q2W=every 2 weeks; Q4W=every 4 weeks

*Week 16 = primary endpoint (HiSCR₅₀ bimekizumab versus placebo)

a Study participant should discontinue from the study from Week 32 on if no partial response is achieved (partial response is defined as $\geq 25\%$ improvement in abscess and inflammatory nodule count relative to Baseline [Week 0] lesion values,)

b Study participants achieving an improvement of at least 25 % in abscess and inflammatory nodule count continue in HS0005 (Extension Study).

c 20-week Safety Follow-up (from last IMP injection) for any study participant who discontinues from study prior to Week 48, or who does not continue in HS0005.

2.3.2 Study periods

2.3.2.1 Screening Period

The Screening Period will last a minimum of 14 days up to a maximum of 5 weeks prior to randomization.

2.3.2.2 Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)

Eligible study participants will be randomized in a 2:2:2:1 ratio as noted in the Study Schema to:

- Bimekizumab 320mg Q2W from Weeks 0 to 48
- Bimekizumab 320mg Q4W from Weeks 0 to 48
- Bimekizumab 320mg Q2W to Week 16, continuing on 320mg Q4W from Weeks 16 to 48
- Placebo to Week 16, continuing on bimekizumab 320mg Q2W from Weeks 16 to 48

2.3.2.3 Safety Follow-up Visit

All study participants not continuing in the extension study, including those withdrawn from IMP as well as those completing all study visits, will have an SFU Visit 20 weeks after their final dose of IMP.

2.3.3 Study duration per participant

The total duration of study participation in HS0003 will be 68 to 71 weeks for those who complete HS0003 and do not participate in the extension study HS0005 and 50 to 53 weeks for those who participate in HS0005 and, thus, do not participate in the 20-week SFU Period. The study is comprised of the following periods:

- Screening Period: 14 days up to a maximum of 5 weeks prior to randomization
- Initial Treatment Period: 16 weeks
- Maintenance Treatment Period: 32 weeks
- Safety Follow-Up Period: 20 weeks after the last dose of IMP

A study participant will be considered to have completed the study if he or she completed the Week 48 visit.

The end of the study is defined as the date of the last scheduled procedure for the last study participant in the study globally, including the SFU, as applicable.

2.3.4 Planned number of participants and sites

A total of approximately 490 study participants will be randomized into the study. The planned number of study sites is approximately 100.

2.3.5 Anticipated regions and countries

The regions planned for study conduct are Western Europe, Central/Eastern Europe, North America and Asia/Australia, with possible extension to other regions and countries.

2.4 Determination of sample size

A total of 490 study participants will be randomly assigned in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to the following treatment arms:

- Bimekizumab 320mg Q2W during Initial Treatment Period (Weeks 0-16) and Maintenance Treatment (Weeks 16-48) Period, N=140
- Bimekizumab 320mg Q2W during Initial Treatment Period (Weeks 0-16), and Bimekizumab 320mg Q4W during Maintenance Treatment Period (Weeks 16-48), N=140
- Bimekizumab 320mg Q4W during Initial Treatment (Weeks 0-16) and Maintenance Treatment Periods (Weeks 16-48), N=140
- Placebo during Initial Treatment Period (Weeks 0-16), and Bimekizumab 320mg Q2W during Maintenance Treatment Period (Weeks 16-48), N=70

The analysis of the primary efficacy endpoint and secondary efficacy endpoints are based on a comparison of bimekizumab versus placebo at Week 16, with alpha adjustment strategy as indicated in Section 4.5.

The power to detect a statistically significant difference for each of the endpoints are shown in Table 2-1. Notably, with a 2-sided significance level of 0.025, the sample size of 140:70 provides 73% power for detecting at least a difference of 1.5 (bimekizumab Q4W vs placebo) for the Worst Skin Pain change from Baseline (CFB) endpoint.

Given the high level of power for each of the primary and key secondary endpoints at the 0.025 significance level for the Q2W comparison (power \geq 0.89), and per the alpha spending strategy, there is a high likelihood that the Q4W comparison of Worst Skin Pain CFB vs placebo will be allowed to be tested against the 0.05 level of significance. The power for this latter test is 81%. The sample size is thus ultimately driven by the Worst Skin Pain CFB endpoint.

Furthermore, the randomization ratio of 2:2:2:1 has been chosen to provide study participants with a high probability (6/7 ~ 86%) of being randomized to active study drug.

After randomization for this study was complete, an additional endpoint to assess Worst Skin Pain response was included in the sequential testing procedure. This additional endpoint is based on the threshold for clinically meaningful change and is defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score at Week 16 among study participants with a score of ≥ 3 at Baseline. Note that the power calculations reported in Table 2-1 for this endpoint are based on the sample size that was initially driven by the Worst Pain CFB endpoint as described above. With a 2-sided significance level of 0.025, the sample size of 104:52 in the subset of participants reporting Baseline HSSDD worst skin pain score at or beyond the threshold for clinically meaningful change (ie, Baseline HSSDD ≥ 3) provides 53% power for

detecting a statistically significant difference between bimekizumab Q4W and placebo in the proportion of Worst Pain responders.

Note that the power at the 0.025 level of significance associated with this endpoint for the comparison between bimekizumab Q2W and placebo is 95%. The Q4W comparison of Worst Skin Pain response vs placebo against the 0.05 level of significance is therefore likely, and the power at this significance level is 65%. Given the strength of the power for the bimekizumab Q2W arm vs placebo across endpoints, it is considered acceptable to have a relatively low power for this final endpoint in the testing sequence for the bimekizumab Q4W treatment arm vs placebo.

Table 2–1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6
Worst Skin Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Skin Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0003 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms

^b Within-participant average of Worst Skin Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Skin Pain score at or above 3 (ie, the threshold for clinically meaningful change from Baseline).

3 DATA ANALYSIS CONSIDERATIONS

3.1 General presentation of summaries and analyses

Statistical analysis and generation of tables, figures, study participants data listings, and statistical output will be performed using SAS Version 9.4 or higher. All tables and listings will use Courier New font size 9.

Descriptive statistics will be displayed to provide an overview of the study results. For continuous variables, descriptive statistics will include number of study participants with available measurements (n), mean, standard deviation (SD), median, minimum, and maximum.

For PRO continuous variables, descriptive statistics will also include variable score, absolute and percentage changes from baseline, Q1 and Q3, 10th, and 90th percentiles.

If no participants have data at a given time point, then only n=0 will be presented. The other descriptive statistics will be left blank. If $n < 3$ then the n, minimum, and maximum only will be presented. The other descriptive statistics will be left blank. If $n = 3$ n, mean, median, minimum and maximum will be presented only. The other descriptive statistics will be left blank.

For categorical variables, the number and percentage of study participants in each category will be presented. Unless otherwise noted, the denominator for percentages will be based on the number of study participants included in the respective analysis set. Study participants with missing data will be accounted for using the following approaches:

- For summaries of demographics and Baseline characteristics: percentages will be based on all study participants in the analysis set and a “Missing” category (corresponding to study participants with missing data for the variable being summarized) will be included as the last row in the list of categories being summarized.
- For summaries of efficacy and safety endpoints, unless otherwise specified: percentages will be based only on those study participants with observed data for the variable being summarized. As the denominator may be different from the number of study participants in the analysis set being considered, the denominator will be displayed in the table. The general format for displaying this will be “n/Nsub (%)”.

Percentages will be presented to 1 decimal place. If the percentage is 100%, a decimal will not be presented. If the count is 0, the percentage will not be presented. Typically, the % sign will be presented in the column header, but not with each individual value.

For the purpose of the tabulations the lower and upper confidence limits for the percentages will be truncated at 0 and 100% respectively. Confidence intervals (CIs) for the response rates in efficacy summaries based on nonresponder imputation (NRI) will be computed using the Wilson approximation.

For bimekizumab PK concentrations, summary statistics will include geometric mean, geometric coefficient of variation (CV), 95% CIs for geometric mean, arithmetic mean, SD, median, minimum, and maximum. All summaries of PK variables will be based on the observed values. No imputation will be used.

Decimal places for descriptive statistics will be subject to the following rules:

- “n” will be an integer
- Mean, SD, and median will use 1 additional decimal place compared to the original data
- CV [%] will be presented with 1 decimal place
- Minimum and maximum will have the same number of decimal places as the original value.

Derived variables in general will display the mean, SD, and median to 1 more decimal place than the variables used in the derivation. If the number of decimal places reported in the raw data is varied then use either the maximum raw number of reported decimal places or 3, whichever is the lowest, as a guide for the descriptive statistics.

When reporting individual values and descriptive statistics for PK concentration data, the following rules will apply with regard to rounding and precision:

- Individual values will be reported to the same level of precision as received from the bioanalytical laboratory
- Descriptive statistics will be reported to the same level of precision as the individual data for the minimum and maximum, and to 1 additional significant figure for the mean (arithmetic and geometric), median, SD, and 95% CI for the geometric mean
- The geometric coefficient of variances (geoCV) will be reported as a percentage to 1 decimal place

Statistical tests of efficacy variables will be presented as 2-sided p-values rounded to 3 decimal places. P-values less than 0.001 will be presented as “<0.001” and p-values greater than 0.999 will be presented as “>0.999.” Statistical comparisons will be 2-sided and will be performed at the 0.05 level of significance unless specified otherwise. The significance levels used as part of the multiple testing procedure are detailed in Section 4.5.

Per protocol, visit windows are ± 3 days from the date of first dose. The 20-week SFU Visit window is ± 7 days from the date of the final dose. All by-visit summaries will contain nominal (ie, scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for assessments that occur within a 3-day time window of a scheduled visit. In that case, the assessment will be mapped to the corresponding scheduled visit and will be used for the analysis. This will only occur for selected vendor data. The only exception to this rule is for unscheduled assessments that occur up to 3 days after the Baseline visit. These unscheduled visits will remain as unscheduled as the Baseline assessment cannot be after the first dose of study drug administration. See Section 3.3 for more details on the definition of Baseline values.

A complete set of data listings containing all documented data as well as calculated data (eg, change from Baseline) will be generated.

3.2 General study level definitions

3.2.1 Relative day

The relative day will be included in different listings and will be calculated as follows:

- If the start (stop) date occurred on or after the first dose, but prior to the double-blind drug stop date, relative day is calculated as start (stop) date minus first dose date + 1
- If the start (stop) date occurred after the last dose of double-blind drug, the relative day to the most recent dose is calculated as start (stop) date minus most recent dose date. The relative day in this situation will be preceded by a ‘+’
- If the start (stop) date occurred before the first dose, the relative day is calculated as start (stop) date minus first dose date. The relative day in this situation will be preceded by a ‘-’.

For AEs, relative days for start and stop dates will be calculated as the number of days since the first injection of the medication. For non-treatment emergent AEs, relative day of onset will be negative if the event started and stopped before the first dose. Relative day will only be computed for fully completed dates and will be missing for partial dates.

Relative day will be calculated from first dose of IMP for all treatment groups, and additionally from first dose of bimekizumab for the Placebo/BKZ 320mg Q2W arm.

3.2.2 Mapping of data from Premature End of Treatment visits

If the Premature end of treatment (PEOT) visit occurs at a scheduled visit as outlined in the schedule of assessments, then no mapping is necessary, and any early withdrawal assessments will correspond to that scheduled visit. The PEOT assessments that occur on a date after a scheduled visit will be assigned to the next scheduled site visit per the protocol following the last visit where assessments were available. This approach means that there is a chance that data will be mapped to a visit where a given assessment was not actually collected per the protocol schedule of assessments. Such data will not be summarized in by-visit tables (though it will be available in the listings).

The only exception to the above rule is for anti-bimekizumab antibody assessments, in which all PEOT assessments will be assigned to the next scheduled visit at which anti-bimekizumab antibodies are assessed. All by-visit summaries will contain nominal visits only. Unscheduled visits will not be mapped to scheduled visits. Note that based on the early withdrawal mapping conventions described above, a mapped PEOT visit is considered as observed at that visit and will be summarized as such in the tables.

3.3 Definition of Baseline values

Section 8.3.3 details the derivation of the Baseline value for the HSSDD assessment. For all other assessments, the below applies.

A Baseline value for a participant is defined as the latest non-missing measurement for that participant up to and including the day of administration of first study medication, unless otherwise stated. If a Baseline assessment is taken on the same day as first administration of study medication, it is eligible to be used as the Baseline value, even in the case that the time of the assessment is recorded as taking place after the time of first study medication administration. This is considered acceptable as this measurement is still the best representation of the Baseline value of the given assessment since it is highly unlikely that the study medication could have an impact on any measurement in such a short period of time. However, such cases should be rare as study center personnel are instructed to do all assessments at the Baseline visit prior to administering study medication. One exception to this rule is plasma concentration of bimekizumab. If Baseline plasma concentration is measured at a time after the first administration of study medication, then it will not be eligible to be considered as a Baseline plasma concentration. Such cases will be discussed with the quantitative clinical pharmacologist.

For randomized participants for whom no start date of treatment is available, the Baseline value will be considered as the last available value on or before the randomization date.

If a Baseline measurement is missing or not collected, and a Screening value is available, the Screening value will be utilized as Baseline instead.

Baseline values for component scores should be computed using components from the same visit where the relevant measurements were recorded prior to dosing. For example, if the Screening visit has all of the components, but the Baseline visit is missing 1 or more components, the Baseline value for the component score should be calculated using the Screening visit values.

When the time of first dose is derived, it will be based on the first injection of study treatment, regardless of whether or not it is an active treatment.

3.4 Protocol deviations

Important protocol deviations are deviations from the protocol that could potentially have a meaningful impact on study conduct or on the primary and key secondary efficacy, key safety, or PK outcomes for an individual participant. The criteria for identifying important protocol deviations will be defined within the appropriate protocol-specific document at study start. Important protocol deviations will be reviewed as part of the ongoing data cleaning process. Important protocol deviations including those that lead to exclusion from the analysis sets will be identified and documented prior to unblinding.

Deviations related to the Coronavirus Disease 2019 (COVID-19) global pandemic are unavoidable deviations from the protocol due to confirmed COVID-19 infection, suspected COVID-19 infection, general circumstances around COVID-19 without infection or any other deviation from the protocol due to COVID-19. COVID-19 protocol deviations will also be reviewed separately as part of the ongoing data cleaning process.

3.5 Analysis sets

3.5.1 Enrolled Set

The Enrolled Set (ES) will consist of all participants who have given informed consent.

3.5.2 Randomized Set

The Randomized Set (RS) will consist of all participants randomized into the study.

3.5.3 Safety Set

The Safety Set (SS) will consist of all participants who received at least 1 dose (full or partial) of IMP. The SS will be used for the demographic, safety, and immunogenicity analyses.

3.5.4 Full Analysis Set

The Full Analysis Set (FAS) will consist of all study participants who received at least 1 dose (full or partial) of IMP and had a valid Baseline measurement and a post-Baseline measurement for abscess, inflammatory nodules, and draining tunnel counts.

3.5.5 Active Medication Set

The Active Medication Set (AMS) will consist of all participants who have received at least 1 dose (full or partial) of bimekizumab. The AMS will be used for summaries of safety that include all data from the Initial Treatment Period and/or Maintenance Treatment Period.

3.5.6 Maintenance Set

The Maintenance Set (MS) will consist of all participants who have received at least 1 dose (full or partial) of bimekizumab in the Maintenance Treatment Period.

3.5.7 Per-Protocol Set

The Per-Protocol Set (PPS) will consist of all study participants in the FAS who had no important protocol deviations affecting the primary efficacy variable. Important protocol

deviations will be predefined and study participants with important protocol deviations will be evaluated during ongoing data cleaning meetings prior to unblinding of the data.

3.5.8 Pharmacokinetics Per-Protocol Set

The Pharmacokinetics Per-Protocol Set (PK-PPS) will consist of study participants who received at least 1 full dose of bimekizumab and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the PK.

3.5.9 COVID-19 Free Set

The COVID-19 Free Set (CFS) will consist of all study participants randomized into the study and who have no COVID-19 impact through Week 16. This analysis set will be used for sensitivity analysis of the primary efficacy endpoint.

3.6 Treatment assignment and treatment groups

It is expected that participants receive treatment as randomized and hence safety analyses will be based on the SS, as randomized. However, if after unblinding it is determined that participants randomized to placebo in the Initial Treatment Period received bimekizumab at any time within the first 16 weeks, then for safety analyses these participants will be reallocated to the appropriate bimekizumab treatment group, unless otherwise specified. Participants randomized to bimekizumab will only be reallocated to the placebo treatment group if they never received bimekizumab. Efficacy analyses will be according to randomized treatment and not actual treatment received.

For the purposes of Initial Treatment Period analyses for the 320mg Q2W dosing regimen, the bimekizumab treatment arms of 320mg Q2W/Q2W and bimekizumab 320mg Q2W/Q4W treatment groups will be pooled.

3.7 Center pooling strategy

Geographic regions have been categorized as North America, Western Europe, Central/Eastern Europe, and Asia/Australia. Below is a table of geographic regions with corresponding countries.

Table 3–1: Geographic regions and corresponding countries

Region	Countries
North America	Canada, United States
Western Europe	Belgium, France, Germany, Italy, Norway, Spain, Switzerland, Denmark, Netherlands
Central/Eastern Europe	Greece
Asia/Australia	Australia, Israel, Turkey

The following center pooling algorithm will be used for each geographic region:

- If a center has 21 or more participants, then no pooling will be done for that center.
- Centers with fewer than 21 participants will be ordered from largest to smallest with pooling proceeding in the following manner:
 - Two or more centers will be combined until the cumulative participant total is at least 21.

- Once a pooled center has at least 21 participants, the process will continue in an iterative fashion for the subsequent centers in the ordered list, where a new pooled center begins each time at least 21 participants has been reached in the previous pool.
- If this iterative process reaches the end of the ordered list of centers where the final pooled center has fewer than 21 participants, then the participants from the centers in that pool will be combined with the pooled center formed in the previous iteration.

This procedure is only to be performed within a geographic region – there will be no pooling of centers across regions.

In the event that the percentage of randomized participants is less than 10% in either of the Asia/Australia or Central/Eastern Europe regions, the two regions will be combined as a geographic region stratum for efficacy modeling, so that there are no modeling convergence issues across efficacy variables.

3.8 Coding dictionaries

Adverse events and medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA®) version 19.0.

Medications will be coded according to the World Health Organization Drug Dictionary (WHODD) version MAR2021 B3 or later. Medical procedures will not be coded.

3.9 Definition of an intercurrent event

Handling of intercurrent events is one of the key elements for the analysis of efficacy endpoints.

An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy (See Section 8.2.2).

Receipt of systemic antibiotic rescue medication is defined as initiating any systemic antibiotic on or after Baseline for any reason (including in response to an AE). The only exception to this rule is if a participant randomized to the antibiotic stratum on a tetracycline antibiotic interrupts their stable dose of tetracycline antibiotic during the study and subsequently restarts the same tetracycline antibiotic as confirmed using the coded preferred term. The restarted dose and frequency of the antibiotic must be the same or lower than the regimen prior to the interruption.

The dates of an intercurrent event are as follows:

- For receipt of systemic antibiotic rescue medication: start date of the antibiotic
- For discontinuation of study treatment due to an AE or lack of efficacy: Last study treatment date + 17 days. Note: study treatment discontinuation includes study discontinuation.

The choice of 17 days is intended to capture the interval between dosing and lesion assessments (14 days), as well as the visit window (3 days).

An additional sensitivity analysis will be conducted where missing data due to COVID-19 will be considered an intercurrent event and will be imputed as a nonresponse at that particular visit. This will be identified when there are missing data at a visit that has been impacted by COVID-19 according the COVID-19 impact CRF page. The date of this intercurrent event will be the date of the impacted visit.

3.10 Changes to protocol-defined analyses

The MS and AMS were added as analysis sets.

The endpoints for PGI-S-HS, PGI-S-SP, PGI-C-HS, and PGI-C-SP were clarified in Section 2.2.1.3 to indicate that absolute change from Baseline will not be calculated, as these are categorical endpoints.

The HiSQOL endpoint was clarified to show that there are only 3 domains: symptoms, psychosocial, activities and adaptations and to add total score.

In Protocol Amendment 4, the secondary efficacy endpoint for skin pain response based on the worst skin pain HSSDD score is defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change at Week 16. In the HS0003 SAP, this endpoint is defined to include the exact value of the clinically meaningful threshold of 3, so that the skin pain response based on the HSSDD worst skin pain score is defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score at Week 16 among study participants with a score of ≥ 3 at Baseline.

The following other efficacy endpoints are included in the protocol but will not be included as part of the analysis:

- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HS Symptom Questionnaire (HSSQ) weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline
- Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48

The calculation of nominal p-values has been added for selected efficacy endpoints. These nominal p-values are not controlled for multiplicity and should not be used to declare statistical significance.

The protocol defines the PK-PPS separately by period, but there will only be one PK-PPS for the overall study.

4 STATISTICAL/ANALYTICAL ISSUES

4.1 Adjustments for covariates

The primary efficacy analyses and selected secondary analyses will be adjusted for the 2 randomization stratification variables:

- Hurley Stage at Baseline (II or III)
- Baseline antibiotic use (Yes or No)

If a participant is stratified in the incorrect stratum (ie, the stratum recorded in the Interactive voice or web Response System differs from the actual stratum the participant belongs to), the actual stratum will be used for the analysis.

The continuous secondary endpoints will also include the Baseline value as a covariate.

The Worst Skin Pain secondary endpoints (change from Baseline continuous endpoint and pain response binary endpoint) will also include analgesic use as a covariate.

4.2 Handling of dropouts or missing data

4.2.1 Efficacy data

Different approaches will be used to handle missing data including how intercurrent events (defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to the given visit) will be considered. A composite strategy will be implemented in which a positive clinical outcome is defined as the study participant achieving HiSCR₅₀ at the given visit and not receiving systemic antibiotic rescue medication, and not discontinuing study treatment due to an AE or lack of efficacy through that visit.

4.2.1.1 Handling missing data for the primary efficacy endpoint

If study participants have an intercurrent event as defined in Section 3.9, then the primary efficacy variable at that timepoint and all subsequent timepoints (whether the data were observed or not) will be set to “nonresponse” as the study participant has not met the criteria for response based on the composite estimand defined in Section 8.2. All remaining missing data for the endpoint will be imputed using multiple imputation Markov-Chain Monte Carlo method (MI-MCMC)/monotone regression for the primary analysis.

In addition, sensitivity analyses using NRI, MI-MCMC/reference-based methods, tipping point analysis, and observed case (OC) methods will be performed, which will assess the impact of different methods of handling missing data.

4.2.1.2 Handling missing data for the secondary efficacy endpoints

For secondary binary efficacy endpoints, intercurrent events will be handled, and missing data will be imputed, using the same methods as for the primary efficacy endpoint. NRI and OC methods will be performed as sensitivity analyses.

For secondary continuous efficacy endpoints, MI-MCMC/monotone regression is the primary method for imputing missing data, regardless of whether the missing data are preceded by an intercurrent event. That is, if an intercurrent event occurs on or before a visit, the result for that visit will be treated as missing and then imputed. If the imputation model cannot converge, last observation carried forward (LOCF) will be used. The OC method will be performed as a sensitivity analysis.

4.2.1.3 Handling missing data for the other efficacy endpoints

For other binary efficacy endpoints, missing data will be imputed using the same method as the primary efficacy endpoint. NRI and OC methods will be performed as sensitivity analyses of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀.

For other continuous efficacy endpoints, the MI-MCMC/monotone regression method will be used to impute missing data as the primary method, regardless of whether the missing data are preceded by an intercurrent event. That is, if an intercurrent event occurs on or before a visit, the result for that visit will be treated as missing and imputed with the missing data. If the imputation model cannot converge, LOCF will be used.

For other ordinal endpoints (EQ-5D-3L, PGI-S-HS, PGI-C-HS, PGI-S-SP, PGI-C-SP), the OC method will be applied as the primary analysis method. No imputation is applied.

4.2.1.4 Missing Data Overview and Summary

In summary, the approaches listed below will be used in this study for handling missing data for efficacy endpoints as appropriate:

- **NRI:** Participants who have missing data at the timepoint of interest are treated as though they did not respond to the treatment. This approach is also referred to as Composite Estimand (NRI).
- **Multiple Imputation (MI) – MCMC / Monotone Regression:** Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using monotone regression.
- **MI-MCMC / Reference-based imputation:** Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using an imputation model based on placebo (reference) data.
- **LOCF:** Post-Baseline missing data are imputed by carrying forward the last available observation (including Baseline).
- **Tipping point analyses:** Assumptions will be made about average outcomes among the subsets of participants who prematurely discontinued study treatment. Various “delta adjustments” will be made to the assumed responses among missing data in each treatment arm with varying degrees of plausibility in order to identify assumptions about the missing data under which the conclusions change (O’Kelly, 2014). Then, the plausibility of such assumptions is discussed.
- **Observed case (OC):** Missing data are not imputed. Only participants with available data who have not discontinued study treatment at the given timepoint are considered. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing.
- **Treatment policy strategy:** All available data observed at the time point of interest will be considered, regardless of the occurrence of intercurrent events. This means the analysis includes on- and off-treatment values collected after study participants prematurely discontinued study treatment but agreed to remain on the study and continued to attend visits and provide assessments at those visits as well as values from study participants who received rescue antibiotic medication. Those observed values will be analyzed according to the study participant’s randomized treatment. Study participants for whom efficacy data cannot be obtained at the week of interest, despite attempts to retain them in the study, will have their data imputed using MI – MCMC / monotone.

The following table depicts which missing data handling approaches will be used based on endpoint priority (primary, secondary, other) and endpoint type (responder, continuous, ordinal).

Table 4–1: Missing data handling approach by endpoint priority and type

End-point Priority	Endpoint Type	Com-posite Estimand (NRI)	Modified Composite Estimand (MI)	MI (MCMC/Reference-based)	Tipping Point	Treat-ment Policy	Hypo- thetical Estimand	OC
Primary	Responder	S ^a	P	S ^a	S	S ^a		S
Secondary included in the statistical testing procedure	Responder	S ^a	P					S
	Continuous						P	S
Secondary not included in statistical testing procedure	Binary	X	X					X
	Continuous						X	X
Other	Responder	X ^d	X					X ^d
	Continuous						X	X ^b
	Ordinal						X ^c	X ^c

MI=multiple imputation, NRI=Nonresponder imputation, OC=Observed case, P=Primary method, S=Sensitivity method, X=Method to be used (no priority designated).

Note: Composite estimand (NRI) refers to the approach in which data preceded by the intercurrent event of study treatment discontinuation due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are also imputed as nonresponse.

Note: Modified Composite Estimand (MI) refers to the approach in which data preceded by the intercurrent event of study treatment discontinuation due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are imputed via a multiple imputation model.

Note: Hypothetical Estimand (MI) refers to the approach where outcomes for study participants without an intercurrent event of study treatment discontinuation are as observed, and outcomes for study participants with the intercurrent event are imputed via a multiple imputation model.

^a Imputation method is applied on continuous data, and responder endpoint is derived from the continuous endpoint based on complete data set where applicable.

^b Required only for by-visit summaries of variables whose value at Week 16 is part of the hierarchical testing procedure.

^c For variables with multiple categories, data will be summarized as observed with an additional missing row to capture missing data at a given visit.

^d NRI/OC sensitivity analysis will be performed only for HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀ summaries.

4.2.2 Missing data algorithms for efficacy analyses

These descriptions focus on the missing data procedures themselves and do not specifically account for dealing with intercurrent events, which is addressed in their respective sections.

4.2.2.1 MI – MCMC / Monotone Regression

In many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied.

Binary endpoint

For a binary endpoint (eg, HiSCR₅₀), the procedure is as follows:

1. Create a data set, one for each treatment group of participants with observed values and those needing estimation by multiple imputation. For the imputation step, a distinction is made between non-monotone missing values (ie, intermittent missing values between completed assessments) and monotone missing values (ie, where all participants data are missing after a given time point).
 - a. For the intermittent missing values, the missing values in each data set will be filled in using the MCMC method with multiple chain, monotone missing data imputing pattern, and non-informative prior for all parameters. Unless specified differently, the first 200 iterations will not be used (the “Burn-in” option). A total of 100 sets of imputations will be performed. The seed used for these imputations will be 762 and all other multiple imputation procedures described in this SAP will use this same seed as well. The resulting 100 imputed data sets will have a monotone missing pattern and will be imputed using a method for monotone missingness. Note that a separate imputation procedure must be invoked in SAS for each treatment group as the seed cannot be set for by groups beyond the first when using a by statement in PROC MI.
 - b. For monotone missing data, monotone regression will be used to impute missing data. A separate regression model is estimated for each variable with missing values (ie, measurement at each time point). Based on the resulting model, a new regression model is then drawn and is used to impute the missing values for the variable. Since the data set has a monotone missing data pattern, the process is repeated sequentially for variables with missing values. This procedure will be based on the 100 imputed datasets generated from the MCMC procedure and will be performed by imputation. The SAS® PROC MI procedure will be used for the imputation.

In both cases, Hurley Stage at Baseline, Baseline antibiotic use, and value of the variable of interest at Baseline and at each post-Baseline visit (prior to the time point of interest) will be included in the imputation model. The post-Baseline values will need to be specified in chronological order in the imputation model so that SAS® PROC MI imputes variables from left to right (eg, the Week 2 value will be first imputed using regression based on the Baseline value, and then Week 4 value will be imputed using regression based on Baseline and Week 2 values, etc). The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model based on the MCMC method will only allow joint multivariate normal variables. Therefore, Hurley Stage at Baseline and Baseline antibiotic use will be re-coded as indicator variables. For Baseline antibiotic use, this will simply be 0 for Baseline antibiotic non-users and 1 for Baseline antibiotic users. For Hurley Stage at Baseline, this will be 0 for Hurley Stage II participants and 1 for Hurley Stage III participants. In order to

achieve model convergence, Baseline antibiotic use may be dropped from the model. If convergence is still not obtained, then Hurley Stage at Baseline may also be dropped from the model. Additionally, if a variable is dropped in order to allow convergence for one model in a study, that variable does not have to be dropped from other models in the study if the model converges without dropping the variable. In other words, model convergence should be evaluated for each efficacy variable independently.

Note: The imputation of each lesion type (inflammatory nodule, abscess, draining tunnel, etc) will be performed separately. The 100 data sets obtained for each type will be merged by imputation number and subject number.

2. For each complete imputed data set, the dichotomous responder variable (eg, HiSCR 0 or 1) will be computed. Each complete imputed data set will then be analyzed based on the logistic regression model.

Note: For derivation of HiSCR response, the AN, inflammatory nodule, abscess, and draining tunnel (fistula/sinus tract) counts at Week 16 in the imputed data sets will be compared directly to the observed Baseline counts to determine response. If values outside of the pre-defined range of values for lesion count (<0) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed draining tunnel (fistula/sinus tract) count of -1 would be changed to 0 before deriving the HiSCR responder variable. Additional ranges for values for secondary and other endpoints are defined in Table 4-2.

Note: Standard rounding rules will also be applied to the imputed values of endpoints that can only take integer values (eg, abscess count). For example, if a study participant has an abscess count imputed as 2.4, this imputed value would be rounded down to 2. This rounding step is performed after the multiple imputation but before deriving the responder variable.

Table 4-2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
Lesion count ^a	0	--	Yes
DLQI total score	0	30	Yes
hs-CRP	LLOQ/2	--	No
HSSDD item score	0	10	No
HSSQ item score	0	10	Yes
HiSQOL symptom status score	0	16	Yes
HiSQOL psychosocial impact score	0	20	Yes
HiSQOL impact on physical activities score	0	32	Yes
EQ-5D-3L VAS	0	100	Yes

Table 4–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
WPAI dimension scores	0	100	No for variables: “Percent work time missed due to problem” and “Percent overall work impairment due to problem”. Yes for variables: “Percent impairment while working due to problem” and “Percent activity impairment due to problem”. These two variables can only take values that are multiples of 10.

^a Lesion counts will be imputed separately for each lesion type (abscesses, draining tunnels [fistulas/sinus tracts], inflammatory nodules, non-draining tunnels [fistulas/sinus tracts], non-inflammatory nodules, HS scars). The imputed lesion counts will be used to derive the endpoints that are dependent on the lesion count data (eg, HiSCR₅₀).

- Estimates of the adjusted responder rate for each treatment group and the associated SE are obtained from the logistic regression of each of the 100 imputed data sets. These estimates will be combined for overall inference using Rubin’s rules, which account for the uncertainty associated with the imputed values (Rubin, 1987), and the combined estimates and SEs will be used to construct 95% CIs using the logit scale. This will be done using SAS PROC MIANALYZE. The combined estimates and 95% CIs on the logit scale will be back-transformed using the inverse logit link function to obtain the adjusted responder rates (%) and associated 95% CIs.

Note: The (unadjusted) proportion of responders will be calculated at each time point by treatment group from the imputed datasets using SAS PROC FREQ. These results will also be combined into an overall inference using SAS PROC MIANALYZE.

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

As the estimates of the odds ratios from the logistic regression models in Step 3 follow a log-normal distribution, a log transformation is needed to normalize these 100 odds ratio estimates. That is because the procedures for combining results from multiple imputed datasets assume that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (Step 3). Additionally, the SE for the odds ratios are transformed as follows:

$$SE = \frac{\log(UCL) - \log(LCL)}{2Z_{\alpha/2}}$$

Where UCL and LCL are the upper and lower confidence limit, respectively, for the CI of the odds ratio from the logistic regression model, and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). The estimates of the log odds ratio for Bimekizumab relative to placebo and the corresponding upper and lower CLs will be provided. The odds ratio will be then estimated by exponentiating the estimate of the log odds ratio. The odds ratio and the confidence limits of the odds ratio will be estimated as follows:

$$OR = \exp(\text{Log odds ratio estimate})$$

$$LCL = OR * \exp(-SE * Z_{\alpha/2})$$

$$UCL = OR * \exp(SE * Z_{\alpha/2})$$

Where OR is the back-transformed estimate of the odds ratio just described, SE is the SE of the log odds ratio and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). These calculations will be done such that odds ratios and corresponding CIs are calculated for the odds ratio of bimekizumab vs. placebo.

Note: If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.

In addition to calculating the odds ratio, associated CIs, and p-values for the pairwise comparisons of bimekizumab and placebo, the estimated proportion of responders (ie, estimated responder rate) and the difference in the proportion of responders between each bimekizumab treatment group and placebo will be estimated, and 2-sided 95% CIs will be created for each difference. The creation of the estimates of the differences will be completed for each bimekizumab treatment group using the process detailed below:

1. Use the logistic regression model to calculate:

Least squares mean estimates of the log odds of bimekizumab (G_B) and placebo (G_P), as well as their corresponding standard errors (S_B and S_P , respectively).

Standard error of the least squares mean estimate of the log odds ratio (S_R)

2. Compute estimates for predicted proportions using the following transformations:

$$P_B = \exp(G_B) / (1 + \exp(G_B))$$

$$P_P = \exp(G_P) / (1 + \exp(G_P))$$

The difference in proportions is then given by:

$$D = P_B - P_P$$

3. Estimate the standard error of D by:

$$S_D = \sqrt{P_B^2(1-P_B)^2S_B^2 + P_P^2(1-P_P)^2S_P^2 + P_B(1-P_B)P_P(1-P_P)S_R^2 - P_B(1-P_B)P_P(1-P_P)(S_B^2 + S_P^2)}$$

The MCMC method for multiple imputation, as previously outlined, will be used to account for missing values. The calculation steps described above will be based on the results provided from the logistic regression model of the multiple imputed datasets. An estimate for the difference in proportions, D, and corresponding standard error, SD, will be computed for each of these datasets. The results from these analyses will be combined into a single estimate of the difference in predicted proportion of response and a 2-sided 95% CI interval using SAS PROC MIANALYZE.

Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. However, in the event there are computational challenges with the imputation model (eg, due to a standard deviation of 0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Continuous endpoint

For continuous endpoints (eg, Change from Baseline in DLQI total score at Week 16), the MI method will be applied as follows:

1. The MCMC/monotone regression method described above in Step 1 for binary endpoints will be performed.
2. Based on the multiply imputed data sets obtained for the given variable, the change from Baseline will be derived for each of the 100 complete imputed data sets based on the observed Baseline value and the observed/imputed post-Baseline values. Note that if the value itself is being summarized, no additional derivation is needed.
3. If a statistical model is being used for the analysis of the variable, then that will be performed for each imputation in this step. If no statistical model is being used, then simple descriptive statistics will be calculated.
4. For data excluding hs-CRP, the following rules apply. The results of the 100 imputed data sets (based on the statistical model or descriptive statistics) are combined with means and standard errors calculated using Rubin's rules (via PROC MIANALYZE). Note that for the calculation of other descriptive statistics such as the median, min, and max, Rubin's rules do not apply. MI estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. Thus, for median, Q1, Q3, minimum and maximum, the following approach will apply:
 - The data will be summarized by treatment, visit, and imputation, and the summary statistics will be computed.
 - Results will be summarized by taking the mean value of each summary statistic at each visit across all imputations.
 - The number of decimal places will remain the same as the original for display purposes (ie, if the mean was presented to 1 decimal place, then the mean of the means will also be presented to 1 decimal place).

For hs-CRP only, the following rules apply. The hs-CRP data will be presented using the geometric mean, 95% CI for the geometric mean, median, Q1, Q3, minimum and maximum. The change from Baseline will be expressed as the ratio to Baseline in the summaries. The following approach will be applied:

- Following the MI procedure, the ratio to Baseline will be calculated for any of the imputed values
- The natural logarithm of the absolute values and of the ratios to Baseline will be calculated

- The logged values will be summarized (using PROC MEANS) by treatment, visit and imputation
- The datasets will be combined using PROC MIANALYZE in order to get the mean and 95% CI estimates from the absolute values and ratios to Baseline (based on logged data) across imputations
- The estimates of the mean and 95% CI will be back-transformed to obtain the geometric mean and 95% CI on the original scale
- For the median, Q1, Q3, minimum and maximum the procedure outlined above for the other endpoints will be followed

If the imputation model cannot converge, LOCF will be used.

4.2.2.2 MI – MCMC / Reference-based imputation

MI-MCMC / Reference-based imputation will be implemented as a supportive analysis for the primary efficacy endpoint (through Week 16).

In this case, placebo will be described as the reference arm.

This procedure will use an imputation model based on data from the placebo group only (Mallinckrodt, 2013). Reference-based MI assumes that the statistical behavior of the bimekizumab and placebo-treated participants after discontinuing study medication becomes that of the placebo-treated participants. All timepoints after discontinuation of the double-blind study treatment for both the bimekizumab and placebo groups will be considered missing. Multiple imputations are used to replace missing outcomes for bimekizumab- and placebo-treated participants who discontinued using multiple draws from the posterior predictive distribution estimated from the placebo arm. For binary efficacy endpoints (eg, HiSCR₅₀ at Week 16), imputation will be done on the lesion counts before assessing the imputed results for HiSCR₅₀ response.

The steps for the procedure are as follows:

1. For non-monotone (intermittent) missing data, MCMC will be used to impute lesion count data, with Baseline antibiotic use, Hurley Stage at Baseline, and lesion count at Baseline and at each post-Baseline visit (in chronological order) being included in the imputation model. This will be done only once for each participant in order to provide a dataset with monotone missing data.
2. Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcome data at visits $t=1, \dots, T$, where T is Week 16 for HiSCR₅₀.
 - a. *Initialization.* Set $t=1$ (Baseline visit)
 - b. *Iteration.* Set $t=t+1$. Create a data set combining records from bimekizumab- and placebo-treated participants with columns for covariates (Hurley Stage at Baseline and Baseline antibiotic use) and outcomes at visits 1 to t . Outcomes for all bimekizumab-treated participants are set to missing at visit t and set to observed or previously imputed values at visits 1 to $t-1$. Outcomes for placebo-treated participants are set to observed at visit t or observed or previously imputed values at visits 1 to $t-1$. The outcomes should be sorted in chronological order in the model.

- c. *Imputation.* Impute missing values for visit t using previous outcomes for visits 1 to t-1, Baseline antibiotic use, and Hurley Stage at Baseline. Note that only placebo data will be used to estimate the imputation model since no outcome is available for bimekizumab-treated participants at visit t. Consequently, the input dataset should include all study participants from placebo but only study participants from the bimekizumab arm that have values at timepoint t missing.
 - d. Repeat steps 2a-2c, 100 times with different seed values (seeds ranging from 853 to 952) to create 100 imputed complete data sets. Study participants whose missing values were imputed in the last PROC MI call will be included in the input dataset for the next PROC MI call. Standard rounding rules will be applied to the imputed values. If the MI procedure yields a value outside of the pre-defined range for the given variable, the value will be updated to be within the predefined range of values for the endpoint of interest. For example, the imputed value for inflammatory nodules will be updated to 0 in the case of an imputed value less than 0.
 - e. *Analysis.* For each completed data set, the outcome (response or change from Baseline) will be calculated using the complete datasets (with no missing data).
3. Each complete imputed data set will then be analyzed based on the statistical model specified in this study (logistic regression). The Week 16 results from logistic regression of each of the 100 imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

4.2.2.3 Tipping Point Analysis

Tipping point analyses will also be implemented as a supportive analysis for the primary efficacy endpoint.

The objective of the tipping point analyses is to evaluate the sensitivity of results to departures from the missing at random assumption and to identify the point at which departures cause results to "tip" from statistically significant to statistically non-significant. As such, these tipping point analyses will only be performed if the primary efficacy analysis results in a statistically significant treatment effect.

For tipping point analyses, data for participants after the intercurrent event date (See Section 3.9) will be changed to missing prior to imputation and, for the bimekizumab treated participants, will be changed to non-response after imputation.

The worst-case scenario will be evaluated first. All missing primary endpoint values for study participants randomized to bimekizumab (where missing values include observations after the intercurrent event date and any other missing values) will be imputed as non-responders, while all missing values for placebo-randomized study participants will be imputed as responders.

While there is little justification for such an approach, it makes the most putative assumption possible against a bimekizumab treatment effect. After applying this imputation approach, a logistic regression model consistent with the one described for the primary analysis will be applied. If the p-value for the odds ratio of bimekizumab versus placebo remains significant, then no further tipping point analyses are needed.

If this analysis based on the worst-case scenario results in a p-value that is not significant (eg, greater than 0.025), then additional tipping point analyses will be performed to identify the point at which results switch or “tip” from significant to non-significant. Note that each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo independently for these analyses. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the p-value in this analysis method will be 0.05 instead of 0.025 throughout for that dose. In the tipping point analysis, a shift parameter or delta adjustment is applied to missing, and subsequently imported primary endpoint values (where missing values include observations after the intercurrent event and any other missing values). These delta implemented on the primary endpoint as follows:

1. Data after intercurrent event date (See Section 3.9) will be set to missing.
2. The same MCMC method described in Section 4.2.2.1 (Step 1a) will be implemented for non-monotone (intermittent) missing pattern values, using the same imputation model. This will be based on 100 sets of imputations.
3. Based on the 100 datasets obtained in Step 2, a monotone regression model will be applied (using the same imputation model as in Step 2) as described in Section 4.2.2.1 (Step 1b). This will be based on 1 imputation.
4. Delta adjustments will be made to imputed lesion count values at Week 16, independently in each treatment group as described below.
5. Delta adjusted imputed values will be truncated so that they are within the range of allowable values for each component.
6. Following the delta adjustments for the lesion counts, HiSCR₅₀ will then be derived based on the delta-adjusted multiply imputed data sets obtained for each component.
7. Each of the 100 imputed datasets will then be analyzed using a logistic regression model with factors of treatment group, Baseline Hurley Stage, and Baseline antibiotic use.
8. The results obtained from the 100 logistic regression analyses in Step 7 will be combined for overall inference using Rubin’s rules, and the results obtained for each shift parameter will be presented in a single table.
9. Steps 4 to 8 will be repeated so that, at each iteration, missing values are adjusted with a larger delta than at the previous iteration. The process will go on until the p-value for the odds ratio between bimekizumab and placebo is no longer statistically significant (eg, ≥ 0.025). The odds ratio, 97.5% CI (or 95% depending on the significance level being used for testing), and p-values obtained for each value of delta will be combined in one single table.

The delta adjustments result in study participants randomized to bimekizumab with missing data having a lower probability of response compared to study participants randomized to placebo with missing data. Since HiSCR₅₀ response is an endpoint for which high lesion counts are associated with a less favorable outcome:

- A positive adjustment is applied to the imputed value for study participants randomized to bimekizumab in order to increase the imputed value and decrease the likelihood of response.
- A negative adjustment is applied to the imputed value for study participants randomized to placebo in order to decrease the imputed value and increase the likelihood of response.

To start, imputed values within each lesion type, will be adjusted by the same value in each treatment arm. This adjustment will be 5% of the observed range within that lesion type. Depending on the results obtained, this adjustment will be multiplied for step 9 above (2 times, 3 times the initial adjustment) until the p-value is no longer statistically significant.

Additionally, study participants randomized to bimekizumab with an intercurrent event should be set to non-response, after applying the delta adjustment outlined in Step 6 above. This ensures study participants randomized to bimekizumab do not have a higher probability of response in the tipping point analyses compared to the primary analysis (ie, a study participant randomized to bimekizumab who is non-responder in the primary analysis cannot become a responder in the tipping point analyses).

4.2.3 Rationale for estimand

Intercurrent events have been identified within the estimands for this study because of their potential to impact efficacy assessments linked with the primary and secondary study objectives. In order to account for the effect of any observed post-randomization intercurrent events on the efficacy analyses, the following estimand strategies will be implemented when evaluating the primary and secondary efficacy endpoints:

- A composite estimand strategy will be used for the primary analysis of the primary and binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, HS worst skin pain response).
- A hypothetical estimand will be used for the primary analysis of the continuous secondary endpoints (CFB in DLQI total score and in “worst skin pain” item for the HSSDD).

4.2.3.1 Composite estimand

A composite estimand strategy as defined in Section 8.2.2 allows incorporation of the two intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) within the definition of the endpoint. These intercurrent events are considered meaningful to the efficacy outcome following receipt of study medication. For example, within the proposed composite estimand framework, a randomized study participant who discontinues from study treatment due to lack of efficacy prior to Week 16 will be considered a treatment failure at Week 16 regardless of the lesion count assessment performed at that visit.

The assumptions and robustness of the primary analysis (modified composite estimand as defined in Section 8.2.2) will be assessed through the sensitivity analyses defined in Section 8.2.3. The impact of intercurrent event handling and data imputation methods on endpoint derivation will also be assessed via the analyses of lesion counts and derived HiSCR variables as specified in Section 8.4.2.1 and Section 8.4.1.1, respectively.

4.2.3.2 Hypothetical estimand

The hypothetical estimand is defined in Section 8.3 and involves a data-driven approach to account for the potential impact of intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) on the analysis of continuous efficacy endpoints. Under this framework, outcomes for study participants without an intercurrent event are analyzed as observed. Conversely, outcomes for study participants with an intercurrent event are imputed via a multiple imputation model, ie any

recorded data on or after the intercurrent event will be set to missing and imputed via multiple imputation following the strategy established in Section 4.2.2.1.

4.2.4 Dates and times

For analyses of AEs and concomitant medication usage, a complete date is required in order to correctly identify the AE or medication as occurring during treatment or not, and for correctly assigning an AE or concomitant medication to the Initial Treatment Period or Maintenance Treatment Period.

For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the participant data listings (ie, no imputed values will be displayed in data listings).

Partial AE and concomitant medication start dates will be imputed as follows:

- Imputation of Partial Start Dates
 - If only the month and year are specified:
 - If the month and year of first dose of study medication is the same as the month and year of the partial start date, then use the date of first dose of study medication,
 - Else, if the month and year of the partial start date are the same as the month and year of a study medication switch date, then use the date of study medication switch,
 - Otherwise, use the 1st of the month of the partial start date;
 - If only the year is specified:
 - If the year of first dose of study medication is the same as the year of the partial start date, then use the date of first dose of study medication,
 - Else, if the year of the partial date is the same as the year of a study medication switch date, then use the date of study medication switch,
 - Otherwise, use the 1st of January of the year of the partial start date;
 - If the start date is completely unknown:
 - If the stop date is unknown or not prior to the date of first dose of study medication, then use the date of first dose of study medication,
 - If the stop date is prior to the date of first dose of study medication, then use the 1st of January of the year of the stop date.
- Imputation of Partial Stop Dates
 - If only the month and year are specified, :
 - Use the last day of the month of the partial stop date;
 - If only the year is specified
 - use December 31st of the year of the partial stop date;
 - If the stop date is completely unknown,

- Do not impute the stop date.

Note that if the stop date or the imputed stop date is prior to the imputed start date, then follow the procedure outlined below:

- If only the year of the start date is specified:
 - If the year of start date is the same as the year of first dose of study medication and the imputed stop date is after the date of first dose of study medication, then set the start date to the date of first dose of study medication,
 - Otherwise, set the 1st January of the year of the start date;
- If only the month and year of start date are specified:
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is on or after the date of first dose of study medication then set the start date to the date of first dose of study medication,
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is before the date of first dose of study medication then set the start date to the 1st of the month of partial start date.

Missing start times for medications will be imputed as 00:00h or with the time of dosing for events occurring on the date of IMP administration in case of missing hour and minute. Otherwise start times with only missing minutes will be imputed with :00 or with the minutes of dosing for events occurring on the date and hour of IMP administration.

In the event of ambiguity or incomplete data that makes it impossible to determine whether a medication was concomitant or an AE was treatment emergent, the medication will be considered as concomitant or the AE will be considered treatment emergent. Similarly, in the event of ambiguity or incomplete data which makes it impossible to determine whether a medication or AE is to be assigned to the Initial Treatment Period or to the Maintenance Treatment Period (or both, for medications), then the medication will be assigned to both Treatment Periods, and the AE will be assigned to the Initial Treatment Period.

4.3 Interim analyses and data monitoring

4.3.1 Data monitoring committee

An independent data monitoring committee (DMC) will periodically review unblinded efficacy and safety data to assess the benefit/risk of bimekizumab in study participants with moderate to severe HS. Efficacy data summaries and individual study participant-level data listings may be provided to the DMC to put the safety review in the context of risk/benefit. Any data to be provided is specified per the DMC charter. [REDACTED]

[REDACTED]

4.3.2 Interim analysis

4.4 Multicenter studies

The center-by-treatment interaction will be tested by adding center and a center-by-treatment interaction term (Section 8.2.3.11). In the model, center will be based on the original centers prior to pooling (Section 3.7). However, if the model is unable to converge due to a low number of participants at a given center, a pooling by center will be applied in order to allow the model to converge. If convergence is still not achieved, a pooling by region will be applied. If convergence still cannot be achieved, this analysis will not be performed. Detailed strategy in Section 3.7 will be applied.

4.5 Multiple comparisons/multiplicity

To control the overall type I error rate at 0.05 for the multiple comparisons in the primary and secondary efficacy endpoints, a closed testing procedure under a parallel gatekeeping framework will be applied (Sun, 2018).

Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$). Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

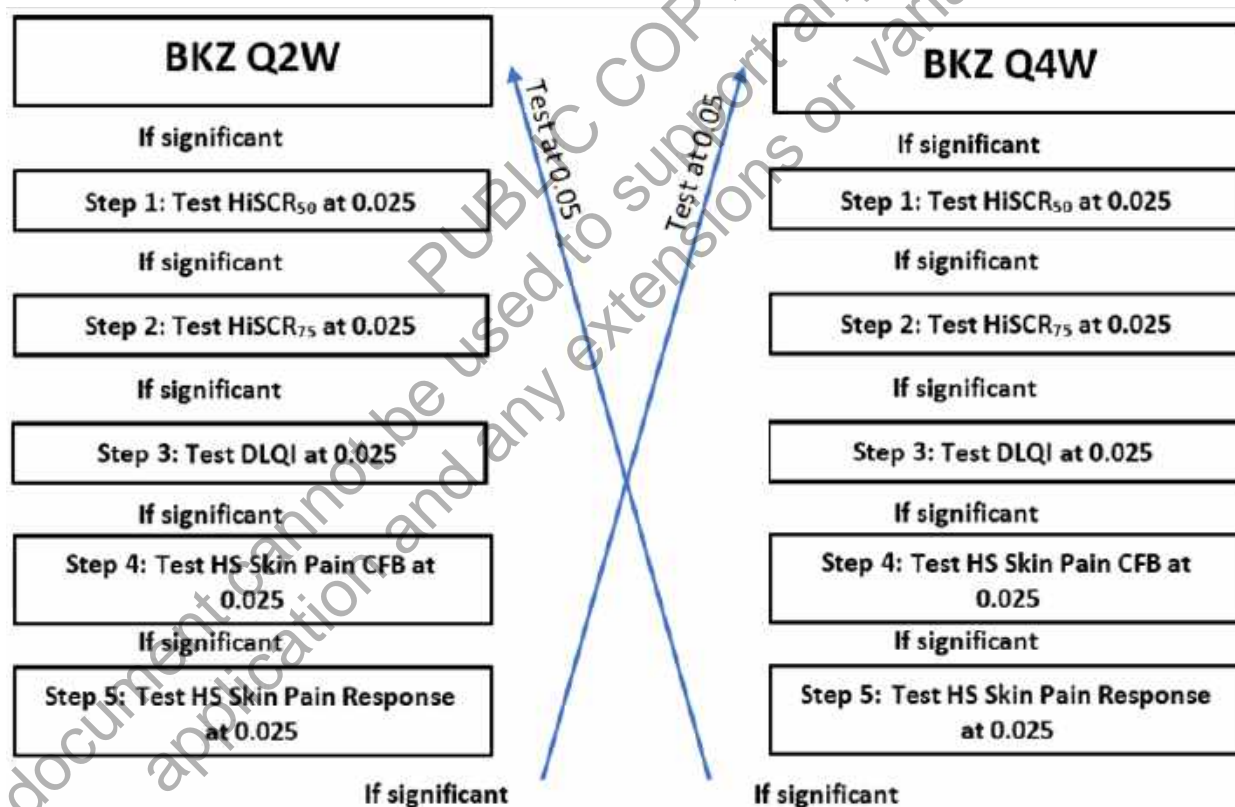
1. Step 1: Test HiSCR₅₀ at significance level 0.025.
2. Steps 2 to 5 – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown in Figure 4-1, moving to the next step only if significance achieved at 0.025.
3. In the event that Step 5 is significant at 0.025 for a given dose, then Steps 1 to 5 will be repeated for the other dose using a significance level of 0.05.

The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment using the analysis methods specified in Section 8.3:

1. Proportion of study participants who achieve HiSCR₇₅ at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo

- b. bimekizumab 320mg Q4W vs placebo
2. Absolute CFB in DLQI Total Score at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
3. Absolute CFB in Skin Pain Score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HSSDD.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
4. Skin pain response at Week 16, based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo

Figure 4-1: Sequence of testing



AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks
HS skin pain response is tested among study participants with a score of ≥ 3 at Baseline.

4.6 Use of an efficacy subset of participants

A sensitivity analysis of the primary endpoint will be performed based on the FAS, the PPS, and the CFS.

4.7 Active-control studies intended to show equivalence

Not applicable.

4.8 Examination of subgroups

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, and worst skin pain response endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period. Additional subgroup analyses will be performed on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, and skin pain response endpoints which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

The following subgroup variables will be determined using Baseline data, except for analgesic use, lesion intervention, and antibody positivity:

- Age (<40 years, 40 to <65 years, ≥65 years)
- Gender (male, female)
- Disease duration (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for analysis.

- Region (North America [Canada, USA], Western Europe [Belgium, France, Germany, Italy, Norway, Spain, Switzerland, Denmark, Netherlands], Central/Eastern Europe [Greece], Asia/Australia [Australia, Israel, Turkey])
- Weight (≤100 kg, >100 kg)
- BMI (<25 kg/m², 25 to <30 kg/m², ≥30 kg/m²)
- Race (Black or African American, White, All Other Races [American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other/Mixed])
- Systemic antibiotic therapy at randomization (yes, no)
- Prior biologic therapy for any indication (yes, no)
- Prior biologic therapy for HS (yes, no)
- Hurley Stage at Baseline (II or III)
- Analgesic users (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period (Section 6.4.2 specifies how participants are classified as analgesic users)
- Lesion intervention (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period

- Antibody positivity (confirmatory assay: negative or positive; see Section 9.3.2)
- Antihistamine users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users) (applicable only to the skin pain response endpoint)

The following subgroups for analysis on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score will be determined based on medication use during the Initial Treatment Period:

- Antihistamines users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users)
- Analgesics users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as analgesic users)
- Systemic antibiotic therapy start/increase after randomization during the Initial Treatment Period (yes, no)

All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

5 STUDY POPULATION CHARACTERISTICS

5.1 Study participant disposition

Summaries of reasons for screen failures (for ES), disposition of participants (for ES), disposition of analysis sets (for RS), disposition and discontinuation reasons in the Initial Treatment Period (for RS) and the Maintenance Treatment Period (for MS), as well as the participants who discontinued due to AEs in the Initial Treatment Period (for RS) and Maintenance Treatment Period (for MS) will be produced. The disposition of participants for all participants screened will include the number of participants included in each analysis set (ES, RS, SS, FAS, AMS, MS, PPS, and PK-PPS) overall and by site.

Participants are defined as completing the Initial Treatment Period if they have a Week 16 visit, or if they fail to attend the Week 16 visit but attend at least one visit in the Maintenance Treatment Period.

The following listings for participant disposition will be provided: participants who did not meet study eligibility criteria (for ES), participant disposition (for ES), participant discontinuation (RS), visit dates (for RS), participant analysis sets (for ES), participants excluded from efficacy analysis (for RS).

To assess participant disposition (entry and periods in the study) during the COVID-19 pandemic, study participant disposition will also be assessed by period of the COVID-19 pandemic (pre – during – post), by comparing the dates of visits (or events) to the dates of the COVID-19 pandemic period. The dates to categorize the periods of the COVID-19 pandemic (pre/during/post) are defined below:

- Pre-COVID-19 pandemic period: Period prior to COVID-19 pandemic start date defined as 11-Mar-2020

- COVID-19 pandemic period: Period from 11-Mar-2020 through the COVID-19 pandemic end date which is currently not defined at the time of approval of the SAP
- Post-COVID-19 pandemic period: Period after the declaration of the end of the pandemic

5.2 Impact of COVID-19

A listing of visits affected by COVID-19 will be presented based on the ES including the visit, date of visit, relationship to COVID-19, impact category and a narrative (short description) of the event. These data will be summarized for non-randomized participants and by treatment group and overall, for enrolled participants.

A summary of study visits by COVID-19 pandemic period (pre – during – post) will be presented for participants enrolled prior to and during the pandemic.

In addition, in order to assess the potential impact of COVID-19 on the collection and reporting of efficacy data, a separate summary on the RS will be presented to display missing data as well as data collected via an alternative modality (e.g.: phone, video call) for efficacy endpoints included in the hierarchy (Section 4.5). For these displays, missing data will be presented only for visits affected by COVID-19, as reported on the dedicated eCRF page. Missing data at other visits and for other reasons will not be included. Note that the remote contingencies for COVID-19 or other exceptional circumstances are not applicable to efficacy assessments and documentation (eg, lesion-based assessments, photography) that require direct face-to-face physician/participant interaction.

5.3 Protocol deviations

Summaries, based on the RS and the MS, displaying the number and percentage of participants with an important protocol deviation (including a summary of participants excluded from the PPS or PK-PPS due to important protocol deviations) by treatment group in the Initial Treatment Period and in the Maintenance Treatment Period, respectively, will be provided. A separate summary of participants with protocol deviations related to COVID-19 will be provided.

A by-participant listing of protocol deviations will be provided. Protocol deviations (eg, missing assessments or visits) related to COVID-19 will be listed separately.

6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

All summaries detailed in this section will be performed on the RS by treatment group. Summaries for demographics and other baseline characteristics will also be repeated on SS and MS. If the RS and SS are identical, the SS summaries will not be created.

6.1 Demographics

Demographic variables will be summarized by treatment group and overall.

The following continuous variables will be summarized using descriptive statistics (number of study participants, mean, SD, minimum, median, and maximum).

- Age (years)
- Height (cm)

- Weight (kg)
- BMI (kg/m²)

BMI (kg/m²) will be calculated as:

$$\text{BMI} = \frac{\text{Weight (kg)}}{(\text{Height (m)})^2}.$$

The following categorical variables will be summarized generally using frequency counts and percentages.

- Age group (≤ 18 , $19 < 65$, ≥ 65 years)
- Age group ($18 < 65$, $65 < 85$, ≥ 85 years)
- Age group (< 40 , $40 < 65$, ≥ 65 years)
- Body weight (≤ 100 kg, > 100 kg)
- Gender
- Race
- Ethnicity
- Ethnic subgroup
- BMI (< 25 kg/m², 25 to < 30 kg/m², ≥ 30 kg/m²)
- Region
- Smoking history
- Country

By-participant listing of demographics for all study participants screened will be provided.

Childbearing potential and lifestyle will be collected at Screening.

6.2 Other Baseline characteristics

The following Baseline disease characteristics will be summarized by treatment group:

- Lesion counts by anatomical region and lesion type, total lesion counts across anatomical regions by lesion type, Hurley Stage by anatomical region and worst overall Hurley Stage across anatomical regions
- IHS4 score, individual items of the HSSDD, HS-Physician's Global Assessment, DLQI total score and HiSQOL domain and total scores
- hs-CRP
- Duration of disease (years)

Duration of disease (years) will be calculated as:

$$\text{Disease Duration} = \frac{(\text{Date of randomization} - \text{Date of HS Diagnosis}^1)}{365.25}.$$

¹If the date of HS diagnosis is partial, it will be imputed to the most recent feasible date (ie, last day of the month if only day is missing, or the last day of the year if day and month are missing). Note that if the date of randomization is missing then the duration of disease will be derived using the date of screening. If the imputed date results in a participant having a disease duration of less than 6 months and the inclusion criterion related to having HS for at least 6 months is confirmed to not have been violated, then the participant's duration of disease will be set to 6 months. If that criterion has been violated, then the participant's duration of disease will be the imputed value of less than 6 months.

- Duration of disease (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for the summary.

- Baseline antibiotic use (yes, no) (According to the randomization strata)
- Baseline antibiotic use (yes, no) (Derived)
- Hurley Stage at Baseline (According to the randomization strata)
- Hurley Stage at Baseline (Derived)

In addition, the following Baseline disease characteristics will be summarized by the derived Baseline Hurley Stage and by the derived Baseline antibiotic use and treatment group for the RS:

- IHS4 score
- "worst skin pain score" and "average skin pain score" in the HSSDD
- HS-Physician's Global Assessment
- DLQI total score
- hs-CRP
- Duration of disease (years)
- Total lesion counts

6.3 Medical history and concomitant diseases

Previous and ongoing medical history will be summarized by treatment groups, system organ class (SOC), and preferred term (PT) using MedDRA[®]. Medical procedures are not coded.

The following listings for medical history and concomitant diseases will be provided: medical history, HS history, concomitant medical procedures, previous and ongoing medical history glossary, previous and ongoing medical history conditions, and procedure history.

6.4 Prior and concomitant medications

Prior medications include any medications that started before the start date of study medication. Concomitant medications are any medication that has a start date on or after the start date of study medication, or any medication that has a start date on or before the last dose of study medication + 28 days (whether placebo or bimekizumab).

Any medication that started before the first dose of IMP and continued after will be classified as both prior and concomitant. Such medications will therefore be counted in the tabulations for both prior and concomitant medication.

Details of imputation methods for missing or partial dates are described in Section 4.2.4.

The number and percentage of participants taking prior medications will be summarized by treatment group, overall and by ATC class, presenting Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT. Prior antibiotic medications will be summarized similarly.

The number and percentage of participants taking concomitant medications will be summarized similarly for the Initial Treatment Period and Maintenance Treatment separately. The number and percentage of participants taking concomitant antibiotic medications, antihistamines, and analgesics will be summarized separately by treatment group, overall, and by ATC class, presenting Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT for the Initial Treatment Period and Maintenance Treatment Period separately.

Separate summaries will be presented for participants taking rescue medication for the Initial Treatment Period and Maintenance Treatment Period separately, identified by a 'yes' response to the 'Is this a rescue medication' question on the electronic case report form (eCRF). This summary will be performed separately for analgesic use and antibiotic use.

Additional summaries for the Initial Treatment Period and Maintenance Treatment Period will be presented for participants taking systemic antibiotic medications that qualify as intercurrent events as described in Section 3.9.

The number and percentage of study participants with concomitant vaccines for COVID-19 will be summarized by treatment group, overall and by World Health Organization Drug Dictionary Standardized Drug Grouping (SDG), presenting SDG subgroup, and preferred term. The SDG subgroup Vaccines for COVID-19 will be used to identify vaccines for COVID-19 using the narrow scope; this subgroup is divided further into separate subgroups which is the level that will be presented. The number of individual occurrences of the vaccine for COVID-19 will also be summarized.

A listing of concomitant vaccines for COVID-19 will be provided.

6.4.1 Assignment of medications to study period

The following rules will be used to assign a concomitant medication to a study period:

- **Initial Treatment Period:** a medication will be assigned to the Initial Treatment Period if it has been taken at least once between the first administration of IMP on Day 1 up to Week 16. This includes medications that started prior to the Initial Treatment Period and those that continued into the Maintenance Treatment Period.
- **Maintenance Treatment Period:** a medication will be assigned to the Maintenance Treatment Period if it has been taken at least once between Week 16 and the final visit. This includes medications that started prior to the Maintenance Treatment Period.

Thus, a medication taken from the time of the first drug administration in the Initial Treatment Period to any timepoint after Week 16 will be assigned to both the Initial Treatment Period and the Maintenance Treatment Period.

Methods for dealing with partial dates are specified in Section 4.2.4.

6.4.2 Classification of participants as analgesic, antihistamine users

If a participant has taken a new analgesic/increased regimen of analgesic, or taken an antihistamine, on 1 or more days (need not be consecutive) in a study period (Initial Treatment Period or Maintenance Treatment Period), then for that period the participant will be classified as an analgesic or antihistamine user, respectively. The period under consideration is to match the period as defined for the HSSDD for the Initial Treatment Period or HSSQ for the Maintenance Treatment Period, based on dates/times of the medications taken.

New analgesic/increased regimen of analgesic use, regardless of indication, is defined as an analgesic medication with start date on or after the first dose of study medication. Stable analgesics (ie, analgesics which were taken already before randomization) will not be included in this category of analgesic user. This classification will be used for selected subgroup analyses.

Antihistamine use is identified by considering the ATC classification. This classification is used for analyzing the Worst Itch endpoint and for selected subgroup analyses, by visit, for the Initial Treatment Period and Maintenance Period as applicable.

Additionally, if a participant has taken a new analgesic/increased regimen of analgesic on 1 or more days (need not be consecutive) prior to the Week 16 visit, then for that week the participant will be classified as an analgesic user. This classification will be used to adjust the formal analysis of the Worst Skin Pain secondary endpoints. If there is a visit date but no HSSDD available at the visit, then the analgesic/antihistamine user status for that week will be derived based on the visit date. If there is no visit available, then the weekly analgesic/antihistamine user status will default to the analgesic/antihistamine status for the overall study period.

7 MEASUREMENTS OF TREATMENT COMPLIANCE

Due to the method of administration of the treatments, compliance will be examined in terms of completed injections.

Treatment compliance will be calculated as:

$$\frac{N_{actual}}{N_{expected}} \times 100\%$$

where N_{actual} is the total number of actual (completed) injections, and $N_{expected}$ is the total number of expected injections. In this study, dosing occurs every 2 weeks from Week 0 to Week 46, where 2 injections are administered at each given visit either with active dosing or placebo injection. It is expected that a participant should complete a total of 48 injections by the end of study. If a participant discontinues early, then the number of expected injections is based on the time of early discontinuation relative to the dosing visits. For example, if a participant discontinues after Week 8 visit and prior to Week 10 visit, the total number of expected injections will be 10.

A summary of percent treatment compliance categorized as <75% and ≥75% will be provided by treatment group for each study period (Initial Treatment Period for the RS, Maintenance Treatment Period for the MS, and the combined Initial and Maintenance Treatment Period for the AMS).

A by-participant listing of treatment compliance will be provided.

8 EFFICACY ANALYSES

All efficacy analyses of primary, secondary, and other variables will be performed on the RS unless otherwise specified. All efficacy summary tables will be displayed by treatment sequence unless otherwise specified. The primary and secondary endpoints, and their components, will also be summarized by the derived Hurley Stage at Baseline (grouping each stage and overall), and treatment sequence and by the derived Baseline antibiotic use (yes/no and overall) and treatment group.

8.1 Lesion count assessment

The primary efficacy endpoint and some of the secondary and other efficacy endpoints discussed in Section 8.3 and Section 8.4 are based on the assessment and/or counts of different types of lesions in the following main anatomical regions at each visit.

- Inguinal (groin)
- Axillary (armpit)
- Chest/breast
- Gluteal
- Abdomen including supra pubic
- Back
- Head
- Neck
- Leg
- Other

These anatomical regions are further classified into the following locations, for the left and right sides of the body, as applicable:

- Inguinal excluding genital and pubic area
- Inguinal including genital and pubic area
- Submammary
- Intermammary
- Chest
- Breast
- Gluteal – Buttocks
- Gluteal – Perianal/Perineal
- Scalp
- Face

All “Other” anatomical regions with lesions present will be specified in free text on the eCRF.

The number of each of the following types of lesions will be recorded in each anatomical region, and then summed across all anatomical regions:

- Abscesses
- Inflammatory nodules
- Non-inflammatory nodules
- Draining tunnels (fistulas/sinus tracts)
- Non-draining tunnels (fistulas/sinus tracts)
- HS scars

If participants undergo lesion interventions as specified in the study protocol, the affected lesions will be counted by the Investigator as permanently present, thus accounting for potential bias due to the intervention. Section 8.4.24 specifies how the intervention data will be presented.

8.2 Primary efficacy endpoint

The primary and sensitivity analyses of HiSCR₅₀ response at Week 16 are summarized in Table 8–1.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

		Estimands for Primary Endpoint			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Primary Objective: To evaluate the efficacy of bimekizumab in study participants with moderate to severe HS					
HiSCR ₅₀	Primary (Section 8.2.2)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (nonresponders).	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.1)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.
HiSCR ₅₀	Sensitivity (Section 8.2.3.2)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – Reference-Based Regression under a missing not at random assumption.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.3)	HiSCR ₅₀ response at Week 16	RS	Composite strategy^a , as for the primary analysis.	A tipping point analysis will be used where various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. The odds ratio versus placebo is based on a logistic regression for each value of delta.
HiSCR ₅₀	Sensitivity (Section 8.2.3.4)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a treatment policy strategy , whereby the data from the Initial Treatment Period are used regardless of whether the intercurrent event occurred.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.5)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.
HiSCR ₅₀	Sensitivity (Section 8.2.3.6)	HiSCR ₅₀ response at Week 16	FAS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.7)	HiSCR ₅₀ response at Week 16	PPS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.8)	HiSCR ₅₀ response at Week 16	CFS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.9)	HiSCR ₅₀ response at Week 16	RS	The same two intercurrent events used for the primary analysis will be used. Any missing data due to COVID-19 will also be considered an intercurrent event. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (nonresponders).	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.10)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a stratified Cochran-Mantel-Haenszel (CMH) test. Missing values not preceded by an intercurrent event will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

CFS=Covid-19 Free Set; CMH=Cochran-Mantel-Haenszel; COVID-19=coronavirus disease 2019; FAS=Full Analysis Set; HiSCR=Hidradenitis Suppurativa Clinical Response; IES=intercurrent event(s) strategy; MCMC=Markov Chain Monte Carlo; MI= multiple imputation; PLS=Population-level summary; Pop=Population; PPS=Per-Protocol Set; RS=Randomized Set

^a The composite estimand strategy will be modified in the tipping point analysis such that participants with intercurrent events will be treated as nonresponders only in the bimekizumab treatment groups.

8.2.1 Derivation of HiSCR₅₀ at Week 16

The following algorithm will be applied to derive HiSCR₅₀ at each visit, based on total lesion counts across anatomical regions for the 3 relevant lesion types recorded as specified above in Section 8.1:

1. Calculate the AN count at each visit as the total number of abscesses plus the total number of inflammatory nodules, across all anatomical regions
2. Calculate the percentage change from Baseline in AN count (%ΔAN) at each visit as

$$100 \times (\text{AN count at post-Baseline visit minus Baseline AN count}) / (\text{Baseline AN count})$$

3. Calculate the change from Baseline in the abscess count by subtracting the Baseline abscess count from the abscess count at each post-Baseline visit
4. Calculate the change from Baseline in the draining tunnel (fistula/sinus tract) count by subtracting the Baseline draining tunnel (fistula/sinus tract) count from the draining tunnel (fistula/sinus tract) count at each post-Baseline visit
5. If the %ΔAN is less than or equal to -50%, and the change from Baseline in the abscess count is zero or negative (ie, non-positive) together with a non-positive change from Baseline in the draining tunnel (fistula/sinus tract) count, then the HiSCR₅₀ will be assigned a value of 1 (ie, HiSCR₅₀ is achieved); otherwise, the HiSCR₅₀ will be assigned a value of 0 (ie, HiSCR₅₀ is not achieved).

In cases where the inflammatory nodule, abscess or draining tunnel (fistula/sinus tract) count is missing and will not allow for the HiSCR₅₀ calculation, the rules for handling missing values in the analysis will be applied (Section 4.2.1 and Section 8.2.3).

The primary efficacy endpoint is attained if the participant has a HiSCR₅₀ of 1 at Week 16.

8.2.2 Primary analysis of the primary efficacy endpoint

The primary endpoint is the HiSCR₅₀ response at Week 16 and corresponding analyses are based on the RS. The primary efficacy analysis will evaluate the composite estimand in the RS as described in Table 8–1. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and completion of study treatment through Week 16 without receiving systemic antibiotic rescue medication or discontinuing IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for the primary efficacy analysis:

1. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
2. Study participant-level outcome=HiSCR₅₀ at Week 16.
3. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. More information is provided in Section 3.9. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, and not discontinuing study treatment due to an AE or lack of efficacy through Week 16. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation as defined in Section 4.2.1. The rationale for this composite estimand is provided in Section 4.2.3.1.

4. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

The statistical hypothesis for the HiSCR₅₀ response at Week 16 is that the conditional odds ratio for the HiSCR₅₀ response in the bimekizumab treatment group relative to the placebo group is equal to 1.

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, and Baseline antibiotic use. The odds ratio versus placebo, p-value (from Wald test), and 97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.

The number and percentage of participants who are HiSCR₅₀ responders at Week 16 will be summarized.

By-participant listings of HiSCR responder endpoints will be provided.

8.2.3 Sensitivity analyses of the primary efficacy endpoint

The following sensitivity analyses for the primary efficacy endpoint will be performed to evaluate the assumptions related to the handling of missing data. Details of the estimands for each analysis are described in Table 8-1.

8.2.3.1 Nonresponse imputation

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (ie, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) will be imputed as nonresponse. That is, participants who experience an intercurrent event will be imputed as nonresponder at the timepoint of the event and all subsequent timepoints (including any recorded data after the event), and all missing data will also be imputed as nonresponse.

The same analysis model as in the primary efficacy analyses will then be used on the imputed data set.

8.2.3.2 MI-MCMC / Reference-based imputation

Deviations from the missing at random pattern will be evaluated using a reference-based MI approach (see Section 4.2.2.2). Intermittent missing data will be imputed using the MCMC method. The remaining monotone missing data will be assumed to follow a missing not at random pattern. These data will be imputed using a reference-based approach in which the MI model is based on data from the placebo group, thereby assuming that monotone missing data follow a trajectory similar to the placebo group.

The same analysis model as in the primary efficacy analyses will then be used on the imputed data set.

8.2.3.3 Tipping point analysis

Tipping point analyses will be performed to evaluate missingness assumptions. Various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. It will include scenarios where study participants who have missing data and are randomized to bimekizumab have a

lower probability of response compared to study participants who have missing data and were randomized to placebo. This includes the worst-case scenario where study participants who have missing data and are randomized to bimekizumab are considered nonresponders, while study participants who have missing data and were randomized to placebo are considered responders. The goal of the tipping point analysis is to systematically vary assumptions until there is no evidence of a treatment effect (if possible). The plausibility of such required delta adjustment will then be assessed.

Refer to Section 4.2.2 for more details on the methodology.

8.2.3.4 Treatment policy

The treatment policy strategy for addressing intercurrent events will be considered. This will be based on an analysis of all available data at Week 16 regardless of the occurrence of intercurrent events. This analysis will use the same models specified for the primary analysis, where study participants are analyzed according to their randomized treatment, even if they discontinued study treatment for any reason prior to Week 16. Even though efforts will be made to collect the primary outcome data for all study participants at Week 16, there may still be some study participants for whom Week 16 efficacy data cannot be obtained. In this case, missing data will be imputed using multiple imputation under the assumption of MAR (see Section 4.2.2). The same analysis model as in the primary efficacy analyses will then be used on the imputed data set and the resulting inferential statistics will then be combined into a single inference using Rubin's rule.

8.2.3.5 Analysis on observed cases

An additional supportive analysis will be based on observed data only for study participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data at Week 16 will be treated as missing (see Section 4.2.2).

The same analysis model as in the primary efficacy analyses will then be used on the imputed data set.

8.2.3.6 Analysis on FAS

The primary efficacy analyses from Section 8.2.2 will be repeated based on the FAS.

8.2.3.7 Analysis on PPS

The primary efficacy analyses from Section 8.2.2 will be repeated based on the PPS.

8.2.3.8 Analysis on CFS

The primary efficacy analyses from Section 8.2.2 will be repeated based on the CFS.

8.2.3.9 Analysis including COVID-19 impact as intercurrent event

An additional sensitivity analysis will include an additional intercurrent event. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and completion of study treatment through Week 16 without receiving systemic antibiotic rescue medication or discontinuing IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for this sensitivity efficacy analysis:

1. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
2. Study participant-level outcome=HiSCR₅₀ at Week 16.
3. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication, discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16, or missing data due to COVID-19. More information is provided in Section 3.9. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, not discontinuing study treatment due to an AE or lack of efficacy through Week 16, and not having missing data due to COVID-19. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation as defined in Section 4.2.1.
4. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

The same imputation techniques and analysis model as in the primary efficacy analyses will then be used.

8.2.3.10 Cochran-Mantel-Haenszel test

The primary efficacy analyses from Section 8.2.2 will be repeated where the CMH test with fixed effects for treatment, Hurley stage at Baseline, and Baseline antibiotic use will be used as stratification variables. Pairwise treatment comparisons will be made based on the CMH test using the p-value for the general association. The odds ratio and associated confidence interval (CI) based on the Wald test will be provided.

8.2.3.11 Center-by-Treatment Interaction

The center-by-treatment interaction will be tested by adding center and a center-by-treatment interaction term in the logistic regression model described in Section 8.2.2. In the model, center will be based on the original centers prior to pooling. However, if the model is unable to converge due to a low number of participants at a given center, a pooling (see Section 3.7) will be described in order to allow the model to converge. In order to obtain reasonable estimates of variability for a treatment arm at a given center, a minimum of 21 participants will be considered acceptable for a center to be included in the model without pooling. Given the 2:2:2:1 randomization allocation scheme, this should provide a minimum of about 12 participants in the bimekizumab 320mg Q2W treatment group, 6 participants in the bimekizumab 320mg Q4W treatment group, and 3 participants in the placebo treatment group. Centers with fewer than 21 participants will be eligible for pooling. The pooling algorithm used is described in Section 3.7.

In order to achieve model convergence, other explanatory variables eg, Baseline Hurley Stage and Baseline antibiotic use may be dropped from the model. If model convergence is still not achieved, region and a region-by-treatment interaction term will be added to the model instead. Regions are defined in Section 3.7.

If the center-by-treatment interaction is not found to be significant ($\alpha=0.05$), then no further analyses will be performed. On the other hand, if the interaction is significant, further analyses

will be conducted to determine which center or centers may be the source of interaction. This will be done by running the logistic regression model (including the interaction term) where each center will be systematically removed from the model. This impact of a given center will be based on the change in the interaction p-value when that center is removed. The center or centers that appear to be driving the significant interaction effect will then be removed from the model to verify that conclusions remain the same with or without the influential center(s). This sensitivity analysis will be based on RS with MI/MCMC Monotone Regression for missing data.

8.3 Secondary efficacy endpoints

The secondary efficacy analyses will be performed based on the RS. Sensitivity analyses of the secondary endpoints will be performed on the CFS.

Missing data handling and sensitivity analyses of the secondary efficacy endpoints are described in Section 4.2.1.2.

The analyses of the secondary endpoints are summarized in Table 8–2 .

Table 8–2: Estimand Details and Attributes for Secondary Analyses

		Estimands for Secondary Endpoints			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Secondary Objective: Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS					
HiSCR ₇₅	Secondary (Section 8.3.1)	HiSCR ₇₅ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
DLQI	Secondary (Section 8.3.2.1)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.
DLQI	Secondary - Sensitivity (Section 8.3.2.2)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a DLQI total score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.3.1)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.
HSSDD	Secondary (Section 8.3.3.2)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.4.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.
HSSDD	Secondary Sensitivity (Section 8.3.4.2.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary Sensitivity (Section 8.3.4.2.2)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

AE=adverse event; ANCOVA=analysis of covariance; DLQI=Dermatology Life Quality Index; HiSCR=Hidradenitis Suppurativa Clinical Response; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; IES=intercurrent event(s) strategy; LSMD=Least Squares Mean Difference; MCMC=Markov Chain Monte Carlo; MI=multiple imputation; PLS=Population-level summary; Pop=Population; RS=Randomized Set

^a Analysis includes all study participants in the RS with a Baseline HSSDD Worst Skin Pain score of 3 or higher.

8.3.1 HiSCR₇₅ at Week 16

A categorical response variable, HiSCR₇₅ at Week 16 is defined to be equal to 1 if %ΔAN is less than or equal to -75%, and the change from Baseline in the abscess count is zero or negative (ie, non-positive) together with a non-positive change from Baseline in the draining tunnel (fistula/sinus tract) count, and 0 otherwise. This definition is introduced for identifying participants who respond to the treatment (1 = responder, 0 = nonresponder). The definition of percentage improvement from Baseline is given in Section 8.2.1.

For HiSCR₇₅ at Week 16, logistic regression as specified for the primary analysis will be implemented to test for superiority. The same analysis approach as outlined for the primary efficacy endpoint will be applied.

8.3.2 Change from Baseline in DLQI Total Score at Week 16

The DLQI is a questionnaire designed for use in adult participants with skin diseases and has been used in patients with HS. This is a validated, quality-of-life questionnaire that covers 6 domains including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment, as assessed over the past week.

The scoring of each answer for the DLQI is as follows:

Table 8–3: Dermatology Life Quality Index

DLQI Scoring	
Response	Score
Very much	3
A lot	2
A little	1
Not at all	0
Not relevant	0
Question unanswered	0
Q7: ‘prevented work or studying’ = yes	3

The DLQI total score is calculated by adding the score of each question. The maximum score is 30, and the minimum score is 0. The higher the score, the more quality of life is impaired.

Meaning of DLQI Total Score

0-1 = no effect at all on patient’s life

2-5 = small effect on patient’s life

6-10 = moderate effect on patient’s life

11-20 = very large effect on patient’s life

21-30 = extremely large effect on patient’s life

This categorization will not be utilized in the analysis.

Because Q7 has a sub-question (referred to as Q7a here) after the leading yes/no question, some clarifying rules for scoring are provided:

- If Q7 is marked as “yes”, a score of 3 is given regardless of the responses to Q7a.
- If Q7 is marked as “no”, “not relevant”, or is missing and Q7a is “A lot”, a score of 2 is given.
- If Q7 is marked as “no”, “not relevant”, or is missing and Q7a is “A little”, a score of 1 is given.
- If Q7 is marked as “no”, “not relevant”, or is missing and Q7a is “Not at all”, a score of 0 is given.
- If Q7 is marked as “no” or “not relevant” and Q7a is missing, a score of 0 is given.
- If Q7 is missing and Q7a is missing, Q7 is considered unanswered (see below for details on how this impacts the DLQI total score).

If 1 question is left unanswered, this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30. If 2 or more questions are left unanswered, the questionnaire is not scored.

Change from Baseline in DLQI total score is defined as Week 16 DLQI total score minus Baseline DLQI total score.

8.3.2.1 Primary analysis of change from Baseline in DLQI Total score at Week 16

Missing data imputation described in Section 4.2.1.2 will be applied.

Change from Baseline in DLQI total score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use and Baseline value as a covariate. The least square mean (LSM), standard error (SE), and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab: the difference between the LSM, the associated 97.5% CI for the contrasts, and the corresponding p-value will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96. Estimand and intercurrent event details are specified in Table 8–2.

8.3.2.2 Sensitivity analysis of change from Baseline in DLQI Total score at Week 16

A sensitivity analysis using the same analysis model as in Section 8.3.2.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

8.3.3 Change from Baseline in Skin Pain score at Week 16, as assessed by the “worst skin pain” item in the HSSDD

The items on the HSSDD assess patients’ perception of the core symptoms of HS experienced in the past 24 hours: pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point numeric rating scale (NRS). Two items assess skin pain: worst skin pain and average skin pain.

Weekly averages will be derived for each of the items of the HSSDD for weeks matching the post-Baseline dosing weeks up to Week 16. A weekly average is the sum of the scored item over the course of the study week divided by the number of days on which the item was completed. The weekly averages will be relative to the respective visit date except for Baseline, which will be anchored to the first dose of study drug. A weekly average will only be calculated if at least 4 non-missing values (not necessarily consecutive) are available. Otherwise, the HSSDD weekly average for the given question will be set to missing.

Baseline will be computed as the average from the first 7 consecutive day period in which there are at least 4 non-missing entries. That is, first consider the first 7 consecutive days prior to the Baseline visit, but not including the Baseline visit day itself. If there are at least 4 non-missing values (not necessarily consecutive), then the Baseline average will be calculated. If there are less than 4 values, the 7 consecutive day period will move one day earlier. If there are at least 4 non-missing values (not necessarily consecutive) in that period, then the Baseline average will be calculated. This will continue until there are at least 4 non-missing values in a 7 consecutive day

period in the 14 days prior to Baseline. If there is no period in which there are at least 4 non-missing entries, then the Baseline value will be set to missing.

Change from Baseline in worst skin pain score is defined as the average Week 16 worst skin pain score minus the Baseline worst skin pain score. Missing data imputation described in Section 4.2.1.2 will be applied to the weekly averages and not to the individual daily PRO data.

8.3.3.1 Primary analysis of change from Baseline in skin pain score at Week 16

Change from Baseline in worst skin pain score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use, analgesic use (Section 6.4.2) and Baseline value as a covariate.

The LSM, SE, and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab, the difference between the LSM, the associated 97.5% CI for the contrasts, and the corresponding p-value will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.

8.3.3.2 Sensitivity analysis of change from Baseline in skin pain score at Week 16

A sensitivity analysis using the same analysis model as in Section 8.3.3.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8-2.

8.3.4 HSSDD skin pain response at Week 16

The analysis set for the analyses of the skin pain response will be restricted to those study participants in the RS with a Baseline worst skin pain score of 3 or higher. The weekly scores and Baseline score are derived as specified in Section 8.3.3.

8.3.4.1 Primary analysis of skin pain response at Week 16

Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, is defined as an improvement in the weekly worst skin pain score of at least 3 points versus Baseline.

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, Baseline antibiotic use, and analgesic use (Section 6.4.2).

The odds ratio versus placebo, p-value (from Wald test), and 97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose. Missing data will be handled as specified in Section 4.2.1.2. Estimand and intercurrent event details are specified in Table 8-2.

The number and percentage of participants who are pain responders at Week 16 will be summarized by treatment group.

By-participant listings of pain response status will be provided.

8.3.4.2 Sensitivity analyses of Skin Pain Response at Week 16

8.3.4.2.1 Nonresponse imputation

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (Table 8–2) will be imputed as nonresponse. That is, participants who experience an intercurrent event will be imputed as nonresponder at the timepoint of the event and all subsequent timepoints (including any recorded data after the event), and all missing data will also be imputed as nonresponse.

The same analysis model as Section 8.3.4.1 will then be used on the imputed data set.

8.3.4.2.2 Analysis on observed cases

An additional supportive analysis will be based on observed data only for study participants with a worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing (see Section 4.2.2).

The same analysis model as in Section 8.3.4.1 will then be used on the imputed data set.

8.4 Other efficacy endpoints

The other efficacy endpoints are listed below and will be evaluated according to the planned assessments in the protocol. This excludes the timepoints for the primary and secondary endpoints specified above in Section 8.2.1 and Section 8.3.

Missing data handling for these endpoints is described in Section 4.2.1.3.

8.4.1 HiSCR endpoints

8.4.1.1 HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀

Categorical response variables HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ are defined to be equal to 1 if %ΔAN is less than or equal to -25%, -50%, -75%, -90%, and -100%, respectively, and the change from Baseline in the abscess count is zero or negative (ie, non-positive) together with a non-positive change from Baseline in the draining tunnel (fistula/sinus tract) count, and 0 otherwise. This definition is introduced for identifying study participants who respond to the treatment (1 = responder, 0 = nonresponder). The definition of percentage improvement from Baseline is given in Section 8.2.1.

HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response will be summarized using frequency tables by treatment group and visit.

A line plot of the HiSCR responder (HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀) rate over time, by treatment group, will be produced.

In order to investigate the effect of intercurrent event handling and missing data handling on the binary response variables, the following iterations of each of the aforementioned plots (HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀, respectively) will be presented, with corresponding summary statistics tables:

- Observed data only (non-imputation for either intercurrent events or missing data)

- Non-response imputation to reflect intercurrent events, non-imputation for missing data
- Full imputation: non-response imputation to reflect intercurrent events, imputation for missing data per Section 4.2.1.3

Bar charts of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ will be produced by visit and Hurley Stage at Baseline for the Initial Treatment Period and the combined Initial and Maintenance Treatment Period. These bar charts will be repeated for HiSCR by visit and Baseline antibiotic use.

Another bar chart of HiSCR rate at Week 16 will be produced by Hurley Stage at Baseline and Baseline antibiotic use.

In addition to the above bar charts, a stacked bar chart displaying whether the criteria are met (yes/no) for each of the component data used in the calculation of HiSCR (%ΔAN, abscess count and draining tunnel [fistula/sinus tract] count) will be generated by treatment group and visit. This graph will summarize the proportion of participants in each of the 8 different yes/no binary responses at each visit (2 x 2 x 2 response combination) at each visit.

8.4.1.2 Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀

See Section 8.4.1.1 for the derivation of HiSCR.

Initial Treatment Period

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first HiSCR_{xx} response, Date of Week 16 visit) – Date of first dose of study medication + 1, here xx represents 25, 50, 75, 90, 100 respectively. All visits including unscheduled visits are considered.

Participants who discontinue study treatment without achieving a given HiSCR response prior to Week 16 visit will be censored at the date of last lesion count assessment. Participants who reach the Week 16 Visit without achieving the given response will be censored at the date of the Week 16 Visit. Participants who experience an intercurrent event prior to achieving a HiSCR response will be censored at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during Initial Treatment Period will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to HiSCR responses will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

Combined Initial and Maintenance Treatment Period

An additional time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the combined Initial and Maintenance Treatment Period will be calculated as above, where the Week 48 visit is considered instead of Week 16.

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the combined Initial and Maintenance Treatment Period will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to HiSCR responses will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.

8.4.1.3 HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response at both Weeks 16 and 48

See Section 8.4.1.1 for the derivation of HiSCR response.

The number and percentage of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at both Weeks 16 and 48 will be summarized based on the RS and MS.

Missing data for the above summaries will be handled using NRI. That is, participants are counted as responders only if they have an observed HiSCR at both Weeks 16 and 48 and have no intercurrent events through Week 48. Otherwise, they are treated as not responding.

8.4.1.4 HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ among Week 16 Responders

See Section 8.4.1.1 for the derivation of HiSCR response.

Summaries of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at each visit from Week 16 through Week 48 will be summarized based on a subset of participants in the MS who achieve response at Week 16. The summaries will be as follows:

- HiSCR₅₀ responder rate based on participants who achieved HiSCR₅₀ response at Week 16
- HiSCR₇₅ responder rate based on participants who achieved HiSCR₇₅ response at Week 16
- HiSCR₉₀ responder rate based on participants who achieved HiSCR₉₀ response at Week 16
- HiSCR₁₀₀ responder rate based on participants who achieved HiSCR₁₀₀ response at Week 16

Line plots of the above HiSCR responder rate categories over time (from Week 16 to Week 48), by treatment group, will be produced.

8.4.1.5 Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders

See Section 8.4.1.1 for the derivation of HiSCR response.

Time to loss of response will be based on the MS and include only participants who had the corresponding HiSCR response at Week 16 (considering intercurrent event handling from the composite estimand described in Section 8.2.2).

Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) is defined as: Date of loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ - Date of Week 16 treatment administration + 1.

Time to loss of response will be estimated and presented using the Kaplan-Meier product-limit method for each treatment. Participants who experience an intercurrent event prior to loss of response will be considered as having lost response on the date of intercurrent event. Participants who reach the Week 48 Visit without loss of response will be censored at the date of the Week 48 Visit. Participants who discontinue treatment or study, for reasons other than those already defined for an intercurrent event, and who have not yet displayed loss of response by the time of withdrawal, will be censored at the date of the last lesion count assessment.

8.4.1.6 Partial response

See Section 8.2.1 for the derivation of AN count.

A partial response is defined as a $\geq 25\%$ reduction in AN count from Baseline (Week 0) at a particular timepoint.

The number and percentage of participants who are partial responders at Week 16 and become HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders in the Maintenance Treatment Period will be summarized by treatment group and visit. These analyses will be based on the subset of participants in the MS that are partial responders but not HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders, respectively, at Week 16. These summaries will be based on observed case data and will not consider the occurrence of intercurrent events.

8.4.2 Lesion count

8.4.2.1 Change from Baseline in lesion count

At each visit, lesion counts will be summarized by treatment group, anatomical region, and lesion type. The following lesion types will be summarized:

- Abscesses
- Inflammatory nodules
- ANs
- Non-inflammatory nodules
- Draining tunnels (fistulas/sinus tracts)
- Non-draining tunnels (fistulas/sinus tracts)
- HS scars

Total lesion counts (ie, the total across all anatomical regions) will be summarized by visit and lesion type, treatment group, overall and by Baseline Hurley Stage. Summaries will also be presented for the change and percentage change from Baseline in lesion counts by anatomical region and total lesion counts by lesion type.

A line plot of the percentage change from Baseline in AN count over time by treatment group will be produced. Separate plots will also be produced for the percentage change from Baseline

in the abscess count, inflammatory nodules count, and draining tunnel count, respectively, over time.

In order to investigate the effect of intercurrent events and missing data handling on lesion count data, the following **additional** iterations of each of the aforementioned plots (percentage change in AN count, abscess count, inflammatory nodule count and draining tunnel count, respectively) will be presented, with corresponding summary statistics tables:

- Intercurrent events:
 - Participants who have experienced intercurrent events
 - Participants who have not experienced intercurrent events
- Missing data handling:
 - Observed lesion counts (i.e., non-imputation)
 - Imputed lesion counts

Lesion count data will be listed by treatment group and anatomical region and will show region-specific Hurley Stage and worst overall Hurley Stage for each participant and visit. The total count for each type of lesion, across all anatomical regions at each visit will be listed separately. For the total abscess count and total draining tunnel (fistula/sinus tract) count, the change from Baseline will be listed; for the AN count, the percentage change from Baseline will be listed.

8.4.2.2 AN count of 0, 1, or 2

The number and percentage of participants with an AN count of 0, 1, or 2 will be presented by treatment group and visit. The denominator for the percentage calculations will be the number of participants in each treatment group in the RS with non-missing data at each visit.

8.4.2.3 AN25, AN50, AN75, AN90, AN100

Categorical response variables AN₂₅, AN₅₀, AN₇₅, AN₉₀, and AN₁₀₀ are defined to be equal to 1 if %ΔAN is less than or equal to -25%, -50%, -75%, -90%, and -100%, respectively, and 0 otherwise. This definition is introduced for identifying participants who respond to the treatment (1 = responder, 0 = nonresponder). The definition of percentage change from Baseline is given in Section 8.2.1. AN25, AN50, AN75, AN90, and AN100 response will be summarized using frequency tables by treatment group for each visit.

8.4.3 Flare by Week 16

See Section 8.2.1 for the derivation of AN count.

Disease flare by Week 16 is defined as at least a 25% increase in AN count with an absolute increase of ≥2 AN relative to Baseline is observed by Week 16. A participant's disease flare status (yes/no) will be determined at each visit.

The number of participants who experience at least 1 disease flare by Week 16 will be analyzed using a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, and Baseline antibiotic use. The odds ratio versus placebo, p-value (from Wald test), and CI will be calculated. Missing data will be handled as described in Section 4.2.1.2.

8.4.4 Flare relative to Baseline

See Section 8.4.3 for the derivation of flare.

Disease flare status will be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in each treatment group. This summary will also include the number of participants with any flare in the Initial Period, Maintenance Period, and the combined Initial and Maintenance Period. A bar chart of percentage of participants with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Initial Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Initial Treatment Period will be presented.

8.4.5 Time to flare by Week 16

See Section 8.4.3 for the derivation of flare.

Time to flare (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first flare, Date of Week 16 visit) – Date of first dose of study medication + 1.
All visits in the Initial Treatment Period including unscheduled visits are considered.

Participants who discontinue study treatment without experiencing a flare prior to Week 16 Visit will be censored at the date of last lesion count assessment. Participants who reach the Week 16 Visit without experiencing a flare will be censored at the date of the Week 16 Visit. Participants who experience an intercurrent event prior to experiencing a flare will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.

The median time to flare, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

8.4.6 Time to flare by Week 48

See Section 8.4.3 for the derivation of flare relative to Baseline.

Time to flare (in days) during the combined Initial and Maintenance Treatment Period will be calculated as:

Min (Date of first flare, Date of Week 48 visit) – Date of first dose of study medication + 1.
All visits in the up to Week 48 including unscheduled visits are considered.

Flare will be defined relative to the Baseline visit. Participants who discontinue study treatment without experiencing a flare prior to Week 48 visit will be censored at the date of last lesion count assessment. Participants who reach the Week 48 Visit without experiencing a flare will be

censored at the date of the Week 48 Visit. Participants who experience an intercurrent event prior to experiencing a flare will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.

The median time to flare, including the 2-sided 95% confidence interval, will be calculated for each treatment.

8.4.7 International Hidradenitis Suppurativa Severity Score System (IHS4)

The IHS4 is a validated scoring tool to dynamically assess HS severity to be used both in real-life and clinical trials settings (Zouboulis et al, 2017). The determination of the IHS4 score requires counting the inflammatory nodules, abscesses and draining tunnels (fistulas/sinus tracts), making it straightforward to apply in both research and clinical practice and easy to use in conjunction with the HiSCR.

$$IHS4 = (number\ of\ inflammatory\ nodules \times 1) + (number\ of\ abscesses \times 2) + (number\ of\ draining\ tunnels\ (fistulas/sinus\ tracts) \times 4)$$

The IHS4 score will be derived based on observed component total lesion count data; in the case of missing component data, the IHS4 score will be missing.

The observed IHS4 score, change and percentage change from Baseline will be summarized by treatment group and visit. Missing IHS4 scores will be imputed using the multiple imputation procedure specified in Section 4.2.2.1, where IHS4 scores will be derived based on the imputed lesion counts.

The IHS4 scores will be categorized into 3 HS categories (mild HS: ≤ 3 , moderate HS: 4-10, severe HS: ≥ 11).

The number and percentage of participants in each category (mild, moderate, severe) will be presented by treatment group and visit. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there are no missing data (OC).

Shift tables for the changes from Baseline in this scale will be presented for each post-Baseline visit by treatment group. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there is no missing data for the change (OC).

8.4.8 HS-Physician's Global Assessment 6-point scale

The HS-Physician's Global Assessment is a validated 6-point scale that is used to measure improvement in inflammatory nodules, abscesses, and draining tunnels. The HS-Physician's Global Assessment scale is defined in Table 8-4.

Table 8–4: HS-Physician's Global Assessment 6-point scale

Score	Rating	Description
0	Clear	No abscesses, no draining tunnels (fistulas/sinus tracts), no nodules
1	Minimal	No abscesses, no draining tunnels (fistulas/sinus tracts), no inflammatory nodules, presence of non-inflammatory nodules
2	Mild	No abscesses or draining tunnels (fistulas/sinus tracts), and less than 5 inflammatory nodules, or Single abscess or draining tunnel (fistula/sinus tract), and no inflammatory nodules
3	Moderate	No abscesses or draining tunnels (fistulas/sinus tracts), and at least 5 inflammatory nodules, or Single abscess or draining tunnel (fistula/sinus tract) in the presence of inflammatory nodules, or Between 2 and 5 abscesses or draining tunnels (fistulas/sinus tracts) with or without inflammatory nodules, up to 10
4	Severe	Between 2 and 5 abscesses and draining tunnels (fistulas/sinus tracts), with inflammatory nodules that are greater than 10
5	Very severe	More than 5 abscesses or draining tunnels (fistulas/sinus tracts)

The gradings will be derived on a participant-level basis (ie, across all anatomical regions) based on the following rules:

- Clear:
 - number of abscesses = 0;
 - number of draining tunnels (fistulas/sinus tracts) = 0;
 - number of inflammatory nodules = 0;
 - number of non-inflammatory nodules = 0;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit;
- Minimal:
 - number of abscesses = 0;
 - number of draining tunnels (fistulas/sinus tracts) = 0;
 - number of inflammatory nodules = 0;
 - number of non-inflammatory nodules ≥ 1 ;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit;
- Mild:

- number of abscesses = 0;
- number of draining tunnels (fistulas/sinus tracts) = 0;
- number of inflammatory nodules ≥ 1 and ≤ 4 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

OR

- sum of number of abscesses and number of draining tunnels (fistulas/sinus tracts) = 1;
- number of inflammatory nodules = 0;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

- Moderate:

- number of abscesses = 0;
- number of draining tunnels (fistulas/sinus tracts) = 0;
- number of inflammatory nodules ≥ 5 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

OR

- sum number of abscesses and number of draining tunnels (fistulas/sinus tracts) = 1;
- number of inflammatory nodules ≥ 1 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

OR

- sum of number of abscesses and number of draining tunnels (fistulas/sinus tracts) ≥ 2 and ≤ 5 ;
- number of inflammatory nodules ≤ 10 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

- Severe:
 - sum of number of abscesses and number of draining tunnels (fistulas/sinus tracts) ≥ 2 and ≤ 5 ;
 - number of inflammatory nodules > 10 ;
 - number of non-inflammatory nodules – no limit;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit;
- Very Severe:
 - sum of number of and number of draining tunnels (fistulas/sinus tracts) > 5 ;
 - number of inflammatory nodule – no limit;
 - number of non-inflammatory nodules – no limit;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit

The number and percentage of participants at each level of the assessment scale (Clear, Minimal, Mild, Moderate, Severe and Very severe) will be presented by treatment group and visit. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there are no missing data for the HS-Physician's Global Assessment.

Shift tables for the changes from Baseline in this scale will be presented for each post-Baseline visit by treatment group. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there is no missing data for the change from Baseline in HS-Physician's Global Assessment.

The HS-Physician's Global Assessment will be listed by treatment group and participant at each visit.

8.4.9 High Sensitivity C-Reactive Protein (hs-CRP)

Concentrations of hs-CRP, changes from Baseline, and percent change from Baseline will be summarized by treatment group and visit, where percent change is calculated as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline hs-CRP} - \text{Baseline hs-CRP}}{\text{Baseline hs-CRP}}$$

Summary statistics will include n, arithmetic mean, SD, median, Q1, Q3, minimum and maximum. For the ratio to Baseline, summary statistics will include n, geometric mean, geoCV, median, Q1, Q3, minimum and maximum.

The ratio to Baseline will be calculated as follows:

$$\text{Ratio to Baseline} = \text{hs-CRP at post-Baseline} / \text{hs-CRP at Baseline visit}$$

The ratio to Baseline will also be summarized by treatment group and visit.

For the hs-CRP data, measurements that are below the limit of quantification (BLQ) will be imputed with half of the lower limit of quantification (LLOQ) for the purpose of calculating summary statistics, changes from Baseline, and ratio to Baseline.

Concentrations of hs-CRP, changes from Baseline, and ratio to Baseline will be listed.

8.4.10 Initiation of systemic antibiotic rescue therapy

See Section 3.9 for the definition of a systemic antibiotic rescue therapy.

The number of participants that use rescue antibiotic therapy will be summarized by treatment group for each period.

8.4.11 Time to initiation of systemic rescue therapy in the Initial Treatment Period

See Section 3.9 for the definition of a systemic antibiotic rescue therapy.

Time to initiation of systemic rescue therapy (in days) during the Initial Treatment Period will be calculated as:

Min (Date of initiation of rescue therapy, Date of change in the dose/type of current antibiotic, Date of Week 16 visit) – Date of first dose of study medication + 1.

Participants who discontinue the study without initiating systemic rescue therapy prior to Week 16 visit will be censored at the date of discontinuation. Participants who reach the Week 16 Visit without initiating systemic rescue therapy will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no Post-Baseline visit.

Time to initiation of systemic rescue therapy will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to initiation of systemic rescue therapy will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to requiring rescue therapy.

The median time to initiation of systemic rescue therapy, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

8.4.12 Time to an intercurrent event in the Initial Treatment Period

See Section 3.9 for the definition of an intercurrent event.

Time to an intercurrent event (in days) during the Initial Treatment Period will be calculated as:

Min (Date of intercurrent event, Date of Week 16 visit) – Date of first dose of study medication + 1.

Participants who discontinue the study without experiencing an intercurrent event prior to Week 16 visit will be censored at the date of discontinuation. That includes participants who discontinue from the study for reasons other than Adverse Event and Lack of Efficacy. Participants who reach the Week 16 Visit without experiencing an intercurrent event will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no Post-Baseline visit.

Time to an intercurrent event will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to initiation of systemic rescue therapy will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to intercurrent event.

The median time to an intercurrent event, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

8.4.13 Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)

See Section 8.3.3 for details on HSSDD Baseline and weekly average definitions and derivations.

Percent change from Baseline in HSSDD responses for worst and average skin pain score is defined as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline HSSDD score} - \text{Baseline HSSDD score}}{\text{Baseline HSSDD score}}$$

Change from Baseline in each HSSDD item (worst skin pain, average skin pain, smell or odor, itch at its worst, and amount of drainage or oozing) score will be summarized using descriptive statistics by treatment group and visit, based on weekly averages. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits. Percentage change will be summarized for the worst and average skin pain items.

Additionally, change from Baseline in each HSSDD item will be evaluated by treatment group at Week 16 via continuous empirical cumulative distribution function (eCDF) plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Change from Baseline in Worst Skin Pain score and Worst Itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

HSSDD response based on clinically meaningful change for the worst skin pain item is defined as at least a 3-point reduction from Baseline in HSSDD among study participants with a score of ≥ 3 at Baseline, based on weekly averages.

The number and percentage of responders based on clinically meaningful change for the worst skin pain item will be summarized by treatment group and visit.

The number and percentage of participants who were responders based on clinically meaningful change at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst skin pain item.

HSSDD response for the worst skin pain and average skin pain items is defined as at least a 30% reduction and at least a 1-point reduction from Baseline among study participants with a score of ≥ 3 at Baseline. The number and percentage of responders for each item will be summarized by treatment group and visit.

The number and percentage of participants who were responders (based on the 30% improvement and 1 point improvement definition) at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst skin pain and average skin pain items.

The number and percentage of participants that complete the HSSDD will be calculated for each visit by treatment group. A participant will be counted as completing the HSSDD at a visit if the minimum number of daily entries is present to calculate the weekly average (see Section 8.3.3). The percentage will be based on the number of participants in the RS. A participant will be considered a completer at a visit if the weekly average can be calculated for that visit.

8.4.14 Hidradenitis Suppurativa Symptom Questionnaire (HSSQ)

The 4 items on the HS Symptom Questionnaire (HSSQ) assesses participants' perception of the core symptoms of HS experienced in the past 7 days - skin pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point NRS.

The change from Baseline score is derived as post Baseline score minus Baseline score. A negative change score indicates a reduction in the score/improvement for the participant.

Summary statistics of the actual values and change and percentage change from Baseline values will be used to summarize each HSSQ item for each visit by treatment group. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Additionally, change from Baseline in each HSSQ item will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

HSSQ response for skin pain item is defined as at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline.

The number and percentage of responders for skin pain item will be summarized by treatment group and visit based on the MS.

The number and percentage of participants who were responders at any timepoint in the Maintenance Treatment Period will be summarized by treatment group for the skin pain score based on the MS.

Change from Baseline in skin pain score and itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

The number and percentage of participants that complete the HSSQ will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the MS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if each of the items are completed at that visit.

8.4.15 DLQI

See Section 8.3.2 for the derivation of DLQI total score.

A DLQI total score of 0 or 1 indicates no impact of the skin disease on health-related quality of life and will be summarized.

A participant is considered to have achieved the minimally clinical important difference (MCID) if their individual improvement (ie, decrease) from Baseline in total score is ≥ 4 . A 4-point improvement in the DLQI total score (DLQI response) has been reported to be meaningful for the participant (within-participant MCID). The summary of MCID will be restricted to participants with a DLQI total score of at least 4 at Baseline to ensure that it is possible for the participant to achieve the MCID.

The DLQI related efficacy variables are defined as follows:

- Change from Baseline in DLQI total score is defined as Post-Baseline DLQI total score minus Baseline DLQI total score.
- Percent of study participants achieving a DLQI total score of 0 or 1 is defined as the number of study participants with DLQI total score of 0 or 1 divided by the number of study participants in RS.
- Percent of study participants achieving a MCID in DLQI total score is defined as the number of study participants with improvement from Baseline in total score of 4 or more divided by the number of study participants in RS that have a Baseline DLQI total score of at least 4.

Missing data for the DLQI total score will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

Change from Baseline in DLQI total score will be summarized using descriptive statistics by treatment group and visit. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Frequency tables will be produced to show the number and percentage of DLQI responders for MCID for each visit by treatment groups.

The number and percentage of participants achieving a DLQI total score of 0 or 1 at each visit will be summarized descriptively using counts and percentages by treatment group and visit.

The number and percentage of participants that complete the DLQI total score will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the total score is calculated at that visit.

A by-participant listing of the DLQI questionnaire, DLQI total score, change from Baseline and DLQI response for MCID and 0 or 1 data will be provided by treatment group.

8.4.16 Hidradenitis Suppurativa Quality of Life (HiSQOL)

The HiSQOL includes 17 items assessed using a 7-day recall period, grouped in 3 subscales: symptoms, psychosocial, activities and adaptation.

The assessment of each item of the HiSQOL is as follows:

Table 8–5: Hidradenitis Suppurative Quality of Life

HiSQOL Scoring	
Response	Score
Unable to do, due to my HS	4
Extremely	4
Very Much	3
Moderately	2
Slightly	1
Not at all	0
I normally do not do this, HS did not influence	0
I am not sexually active	0
I do not work or study	0
Unanswered	0

The HiSQOL total score is calculated by adding the score of each question. The maximum score is 68, and the minimum score is 0.

Subscale scores will be summarized for each of the 3 subscales. The maximum scores for the subscales are 16 (symptoms), 20 (psychosocial), and 32 (activities and adaptations), and the minimum score is 0 for all subscales.

For all scores, the higher the score, the more quality of life is impaired.

Summary statistics of the actual values and change from Baseline values will be used to summarize HiSQOL domain and total scores for each visit by treatment group. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Additionally, change from Baseline in each HiSQOL subscale will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3. The imputed HiSQOL total score will be derived based on the imputed subscales.

The number and percentage of participants that complete the HiSQOL will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the total score is calculated at that visit.

A by-participant listing of the HiSQOL questionnaire, HiSQOL responses, domain and total scores and change from Baseline will be provided.

8.4.17 Patient Global Impression of HS Severity (PGI-S-HS)

The PGI-S-HS is a single item to assess study participants' perceptions of the overall severity of HS over the past 7 days (none, mild, moderate, severe, very severe).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-S-HS will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.18 Patient Global Impression of Change in HS Severity (PGI-C-HS)

The PGI-C-HS is a single item to assess study participants' perception of the change in HS since they started taking the study medication (much better, a little better, no change, a little worse, much worse).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-C-HS will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.19 Patient Global Impression of Severity of Skin Pain (PGI-S-SP)

The PGI-S-SP is a single item to assess study participants' perceptions of the severity of their skin pain from their HS lesions, over the past 7 days (none, mild, moderate, severe, very severe).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-S-SP will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.20 Patient Global Impression of Change in Severity of Skin Pain (PGI-C-SP)

The PGI-C-SP is a single item to assess study participants' perceptions of change in their skin pain from their HS lesions, since they started taking the study medication (much better, a little better, no change, a little worse, much worse).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-C-SP will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.21 Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)

The EQ-5D-3L comprises 5 questions on mobility, self-care, pain, usual activities, and psychological status with 3 possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). In addition, there is a VAS to indicate the general health status with 100 indicating the best health status.

Change from Baseline in EQ-5D-3L VAS scores is defined as Post-Baseline EQ-5D-3L VAS score minus Baseline EQ-5D-3L VAS score.

Responses to EQ-5D-3L will be summarized based on OC only as primary analysis. No imputation is applied to responses to EQ-5D-3L but is applied to EQ-5D-3L VAS scores.

Changes from Baseline in EQ-5D-3L VAS will be summarized using descriptive statistics by treatment group and visit. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

The number and percentage of participants with each response in the EQ-5D-3L will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the EQ-5D-3L will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if each of the domains and VAS are completed at that visit.

8.4.22 Work Productivity and Activity Impairment Questionnaire–Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores

The WPAI-SHP V2.0 is a patient-reported questionnaire that assesses study participant's employment status, work absenteeism, work impairment while working (presenteeism), overall work, and daily activity impairment attributable to a specific health problem. It has been used in several clinical studies of biologic therapy in participants with plaque PSO.

Five out of 6 items of the WPAI-SHP are regrouped into the 4 dimensions, with scores expressed as percentage, where higher numbers indicate greater impairment and less productivity, ie, worse outcomes, as described in the WPAI-SHP scoring rules.

The scoring rules for the WPAI-SHP are as follows:

Questions:

- 1 = currently employed
- 2 = hours missed due to specified problem
- 3 = hours missed other reasons
- 4 = hours actually worked

- 5 = degree problem affected productivity while working
- 6 = degree problem affected regular activities

Scores:

- Percent work time missed due to problem: $[\text{Q2 hours}/(\text{Q2 hours} + \text{Q4 hours})] * 100$
- Percent impairment while working due to problem: $[\text{Q5 score}/10] * 100$
- Percent overall work impairment due to problem: $[\text{Q2 hours}/(\text{Q2 hours} + \text{Q4 hours}) + [(1 - (\text{Q2 hours}/(\text{Q2 hours} + \text{Q4 hours})) \times (\text{Q5 score}/10)]] * 100$
- Percent activity impairment due to problem: $[\text{Q6 score}/10] * 100$

A negative number will indicate a reduction in the score/improvement for participants.

The change from Baseline score is derived as post Baseline score minus Baseline score. A negative change score indicates a reduction in the score/improvement for the participant.

Summary statistics of the actual values and change from Baseline values will be used to summarize WPAI-SHP for each visit by treatment group. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

The number and percentage of participants that complete the WPAI-SHP will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the percentages in each dimension are calculated at that visit.

A by-participant listing of the WPAI-SHP questionnaire, WPAI-SHP domains and change from Baseline will be provided.

8.4.23 Treatment Satisfaction Questionnaire – Medication-9

The TSQM-9 is an abbreviated 9-item version of the TSQM, excluding the side effects of medication domain. The domains included in the TSQM-9 include effectiveness (3 items), convenience (3 items) and global satisfaction (3 items). The TSQM-9 domain scores range from 0 to 100 with higher scores representing higher satisfaction.

The scores for each measure are as follows:

- Global Satisfaction:
 - If no items are missing: $([\text{Sum}(\text{Item 7 to Item 9}) - 3]/14) * 100$
 - If either Item 7 or 8 is missing: $([\text{Sum}(\text{the two completed items})] - 2)/10 * 100$
 - If Item 9 is missing: $([\text{Sum}(\text{Item 7 and Item 8})] - 2)/8 * 100$
- Effectiveness
 - If no items are missing: $([\text{Item 1} + \text{Item 2} + \text{Item 3}] - 3)/18 * 100$

- If one item is missing: $([(\text{Sum}(\text{the two completed items}) - 2]/12) * 100$
- Convenience
 - If no items are missing: $([\text{Sum}(\text{Item 4 to Item 6}) - 3]/18) * 100$
 - If one item is missing: $([\text{Sum}(\text{the two completed items}) - 2]/12) * 100$

Frequency tables will be produced to summarize answers provided to each of the 9 items of the TSQM-9 at Weeks 16 and 48 by treatment group. Responses to TSQM-9 will be summarized based on OC. No imputation will be applied.

The number and percentage of participants that complete the TSQM-9 will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if each of the domains are completed at that visit.

A by-participant listing of TSQM-9 will be provided.

8.4.24 Lesion intervention

Investigators will have the option to perform interventions in the event an acutely painful lesion occurs that requires immediate intervention.

The following rules will be used to assign a lesion intervention to a study period:

- **Initial Treatment Period:** a lesion intervention will be assigned to the Initial Treatment Period if it has been performed between the first administration of IMP on Day 1 up to and including Week 16.
- **Maintenance Treatment Period:** a lesion intervention will be assigned to the Maintenance Treatment Period if it has been performed between Week 16 through the Week 48 visit.

Methods for dealing with partial dates are specified in Section 4.2.4.

A listing of participants who receive any lesion intervention will be provided.

The number and percentage of participants who receive at least 1 lesion intervention will be summarized by treatment group for the Initial Treatment Period, Maintenance Treatment Period, and the combined Initial and Maintenance Treatment Period.

The number and percentage of participants with 2, 3, and 4 or more lesion interventions will also be summarized by treatment group for the Initial Treatment Period, Maintenance Treatment Period, and the combined Initial and Maintenance Treatment Period.

The number and percentage of participants with 2 or more lesion interventions performed on the same lesion will be summarized by treatment group for the Initial Treatment Period, Maintenance Treatment Period, and the combined Initial and Maintenance Treatment Period.

8.5 Additional statistical analyses of other efficacy endpoints

For selected other efficacy variables, it is of interest to perform statistical tests and to calculate inferential statistics. As these tests are not part of the multiplicity-controlled procedure, the associated p-values are considered nominal and are not controlled for multiplicity.

For binary variables, the analysis will follow what was specified for the primary analysis of the primary endpoint and the corresponding p-value reported. Missing values will be imputed as for the primary analysis. For continuous variables, the MI – MCMC / Monotone Regression approach used for the secondary continuous endpoint will be applied for the imputation model. The analysis model will be as for the corresponding secondary continuous endpoint analysis, unless otherwise indicated.

Below is a list of variables for which these nominal p-values will be calculated. The results of these inferential tests will be presented in a single table summarizing the testing performed outside of the multiplicity-controlled testing procedure.

All tests will be for both 320mg Q2W vs Placebo and 320mg Q4W vs Placebo (tested separately) and will be performed for the Week 12 and Week 16 visits only. If the Week 16 visit test is already part of the controlled testing procedure in the primary or secondary analyses, only Week 12 is indicated here.

- HiSCR₅₀ at Week 12
- HiSCR₇₅ at Week 12
- HiSCR₉₀
- HiSCR₁₀₀
- Flare by Week 12
- Flare by Week 16
- Time to flare by Week 12 (based on time-to-event analysis per Section 8.4.5 and adjusted appropriately)
- IHS4 change from Baseline
- IHS4 percentage change from Baseline
- HS Physician's Global Assessment: rate of participants who are Clear or Mild
- DLQI total score change from Baseline at Week 12
- Worst Skin Pain per HSSDD change from Baseline at Week 12
- Skin Pain response per HSSDD at Week 12

9 PHARMACOKINETICS AND PHARMACODYNAMICS

9.1 Pharmacokinetics

Pharmacokinetic variables will be analyzed for all participants in the PK-PPS. Bimekizumab plasma concentrations will be summarized for each treatment at each scheduled visit.

PK summaries will be based on observed values. No imputation will be used. However, if plasma concentration measurements are below the level of quantification (BLQ), then for calculation of the derived statistics the result will be set to ½ of the lower level of quantification (LLOQ). Descriptive statistics including geometric mean, geometric coefficient of variation, and geometric mean 95% CI if applicable will be calculated if at least 2/3 of the values of interest are above the LLOQ. If this is not the case, only median, minimum, and maximum will be presented.

Geometric mean plasma concentration will be plotted by treatment group, and by cumulative antibody status for participants randomized to bimekizumab on linear and log linear scale. In addition, HiSCR₅₀ response (0=not achieved, 1=achieved) assessed at each PK visit (excluding Week 1) will be plotted against the visit's bimekizumab plasma concentration, separated by treatment group. Spaghetti plots of bimekizumab plasma concentrations by week from bimekizumab first dosing separated by treatment group and antibody status will be presented for participants with and without HiSCR₅₀ response at Week 16.

If the dosing for a visit is +/- 7 days out of window, then the plasma concentration from that visit and all subsequent visits will be excluded from the PK summary. In addition, if the PK sampling date is >1 day after the dosing date, then the plasma concentration from that visit will be excluded from the PK summary.

All PK concentrations collected will be listed irrespective of the dosing or sampling occurring out of window.

9.2 Pharmacodynamics

Not applicable.

9.3 Immunogenicity

9.3.1 Autoantibodies

Not applicable.

9.3.2 Anti-bimekizumab antibodies

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-tiered assay approach: screening assay, confirmatory assay, and titration assay. Samples confirmed as positive within the confirmatory assay will be further evaluated for the presence of neutralizing anti-bimekizumab antibodies specific to IL-17AA, IL-17FF or both. Samples will be taken at Baseline, then at study Weeks 4, 8, 12, 16, 20, 24, 36 and 48, and at PEOT and SFU timepoints.

ADAb samples are not analyzed when study participants are on a treatment other than bimekizumab. For study participants who switch from placebo to bimekizumab, samples are analyzed starting at the visit when the switch to bimekizumab occurs (Week 16). The sample at Week 16 will act as the Baseline for that treatment group.

The screening cut point will be used to determine the status of anti-bimekizumab antibodies in the test sample as Positive Screen (PS) or Negative Screen (NS). For samples presenting anti-bimekizumab antibody levels that are PS, a further confirmatory assay will be performed, and the result of which will be reported as either Positive Immunodepletion (PI) or Negative Immunodepletion (NI).

ADAb status for each sample will be derived as follows:

- Sample values that are either NS, or PS and NI and where the bimekizumab concentration is less than the validated ADAb assay drug tolerance limit will be defined as anti-bimekizumab antibody negative.
- Sample values that are either NS, or PS and NI and where the bimekizumab concentration exceeds the validated ADAb assay drug tolerance limit will be defined as inconclusive.

- Sample values that are PS and PI will be defined as ADAb positive (regardless of availability of a titer value)
- Missing or non-evaluable samples will be defined as missing

Positive immunodepletion samples will be titrated, and the ADAb titer (reciprocal dilution factor including minimum required dilution) will be reported. Subsequently, PI samples will also be subject to a neutralizing assay to evaluate the potential of ADAb to neutralize the target binding of bimekizumab (IL-17AA or IL-17FF or both) in vitro.

Cumulative ADAb status will be derived as follows:

The ADAb status (positive, negative or missing) will be considered in a cumulative manner at each time point.

A study participant will be counted positive from the first visit at which the study participant achieved a positive ADAb sample result to the end of the treatment period, regardless of any missing/inconclusive or negative ADAb sample result.

If a study participant has only negative ADAb samples or only one missing/inconclusive sample with all other ADAb samples being negative, the study participant will be classified as negative. An exception remains for the Baseline Visit where only one sample could be available. If the sample is missing/inconclusive, then the sample will be classified as being negative for the cumulative ADAb status.

Otherwise, the study participant will be classified in the missing ADAb category.

Overall ADAb status will be derived as follows:

A study participant will be classified as:

- Positive if the study participant has at least one positive sample up to the time point of interest (regardless of having missing/inconclusive data).
- Negative if the study participant has all the samples negative or only one missing/inconclusive sample with negative ADAb samples up to the timepoint of interest.
- Missing if the study participant has more than one missing ADAb result (or have more than one inconclusive sample) and all other available ADAb samples are negative up to the time point of interest.

ADAb categories will be derived as follows:

- **Pre ADAb negative – treatment-emergent ADAb negative (Category 1):** Includes study participants who are anti-bimekizumab antibody negative at Baseline and anti-bimekizumab antibody negative at all sampling points during the period of interest (one post-Baseline missing/inconclusive sample is allowed for subjects with pre- anti-bimekizumab antibody negative sample). This group also includes study participants who have a missing or inconclusive sample (either missing or inconclusive or insufficient volume) at Baseline (ie, pre-treatment) with all post-Baseline samples as ADAb negative.
- **Pre ADAb negative – treatment-emergent ADAb positive (Category 2):** Includes study participants who are ADAb negative at Baseline and ADAb positive at any sampling points post-Baseline during the period of interest. This group also includes study participants who

have a missing sample (either missing or insufficient volume) at Baseline (ie, pre-treatment) with 1 or more post-Baseline samples as ADAb positive.

- **Pre ADAb positive – treatment-emergent reduced ADAb (Category 3):** Includes study participants who are ADAb positive at Baseline, and ADAb negative at all sampling points post-Baseline during the period of interest.
- **Pre ADAb positive – treatment-emergent unaffected ADAb positive (Category 4):** Includes study participants who are ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with titer values of the same magnitude as Baseline (ie, less than a predefined fold difference from the Baseline titer).
 - For this analysis, this is set at an increase of less than the validated Minimum Significant Ratio (MSR) of 2.07-fold from Baseline.
- **Pre ADAb positive – treatment-emergent ADAb boosted positive (Category 5):** Includes study participants who ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with increased titer values compared to Baseline (equal to or greater than a predefined fold difference increase from Baseline titer which will be defined within the validation of the assay).
 - For this analysis, this is set at an increase equal to or greater than the validated MSR of 2.07-fold from Baseline.
 - Note: for any study participant who is ADAb positive at Baseline and ADAb positive at a post-Baseline time point during the period of interest, but for whom titers are not available to determine treatment unaffected or treatment boosted status, the study participant will be considered as treatment boosted, assuming no other samples are available.
- **ADAb Inconclusive (Category 6):** Includes study participants who have an ADAb positive Baseline (pre-treatment) sample and some post-Baseline samples during the period of interest are missing or inconclusive, while other post-Baseline samples are ADAb negative.
- **Total treatment-emergent ADAb positivity (Category 7 [Categories 2 and 5 combined]):** Includes study participants who are pre ADAb negative – treatment-emergent ADAb positive (Category 2) and pre ADAb positive – treatment boosted ADAb positive (Category 5).
- **Total prevalence of pre- ADAb positivity (Category 8 [Categories 3, 4, 5 and 6 combined]):** Study participants that are tested ADAb positive at Baseline.
- **Missing:** Includes study participants who are ADAb negative, missing, or inconclusive at Baseline with some post-Baseline samples as missing or inconclusive, and other samples as ADAb negative.

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, the following categories can also be used:

- **ADAb positive** – This is defined as study participants who are anti-bimekizumab antibody positive on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- **ADAb negative** – Study participants for who either:

- All samples (including Baseline) are ADAb negative and there are no missing or inconclusive samples
- Only 1 sample is ADAb positive and all other samples (including Baseline) are ADAb negative or missing or inconclusive
- Only 1 sample is missing or inconclusive and the remaining ADAb samples are negative.
- ADAb missing - Defined as study participants who do not fulfil the criteria for one of the 2 groups listed above.

The rationale for requiring at least 2 time points in which ADAb levels are above the specified cut point is to exclude those study participants who have only one occurrence of ADAb levels during the course of treatment. Including such study participants would increase the number of ADAb positive study participants with potentially no impact on efficacy.

In the case that a sample is collected 1 or more days following the scheduled visit date in which the drug was administered, the ADAb results for that sample will be associated with the scheduled visit and summarized accordingly. Such samples will also be considered when ADAb results are summarized over a given study period.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by participants. All analyses will be run on the AMS, unless specified otherwise.

- Summary of ADAb status overall and by each visit separated by treatment group
- Summary of the time-point of the first occurrence of ADAb positivity during the treatment period by treatment group. A plot of the titer by time to first ADAb positivity will be prepared.
- All individual participant-level ADAb results will be listed.
- The number and percentage of participants in each of the 8 ADAb categories during the treatment period by treatment group.
- The prevalence of immunogenicity, separated by treatment group, and defined subcategory, will be reported by visit, defined as (cumulative) proportion of participants having confirmed positive ADAb samples at any visit up to and including that visit. Missing samples will not be included in the denominator.
- The time to achieving treatment-emergent ADAb positivity, separated by treatment group and defined subcategory, will be analyzed based on Kaplan-Meier methods. This will be shown only for Categories 2 and 8 above. Participants will be considered to have an event at the time point at which treatment emergent ADAb positive is first achieved (taking the MSR into consideration for sub-category 5). Participants classified as treatment-emergent ADAb negative will be censored at the time of the last available ADAb result.
- A summary of HiSCR₅₀ responders at Week 16, separated by treatment group, as a function of ADAb titer will be presented graphically.
- Individual plots of plasma bimekizumab concentrations/ ADAb titer both plotted on the Y-axes by visit (x-axis) for the full treatment period (excluding SFU for interim analyses and

including SFU for final analyses) will be presented for participants with and without HiSCR₅₀ response at Week 16.

- Spaghetti plots of ADA_b titer (y-axis) by visit (x-axis), separated by treatment group for all ADA_b positive participants, including Baseline positive participants.
- Box plots of ADA_b titer (logscale) by time to first ADA_b positivity by treatment group.

The groups for defining ADA_b status for safety subgroup analyses are as follows:

- AEs starting before first ADA_b positive result
- AEs starting on or after first ADA_b positive result
- AEs for participants who were always ADA_b negative

This is further explained in Section 10.2.2.

10 SAFETY ANALYSES

All analysis of safety variables will be performed using the SS, MS, and AMS.

The AMS will be used for summaries of safety that include data from the Initial Treatment Period and Maintenance Treatment Period.

Summaries of safety will be presented for the Initial Treatment Period, Maintenance Treatment Period, and combined Initial and Maintenance Treatment Period unless specified otherwise.

10.1 Extent of exposure

Summaries for exposure will be provided. This consists of a descriptive summary of study medication duration in days. In addition, total study medication duration and time at risk will be summarized in years by treatment group and treatment period (ie, the Initial Treatment Period, the Maintenance Treatment Period, and the Initial and Maintenance Treatment Period). Summary of exposure in Maintenance Treatment Period will be on MS. The cumulative study medication duration will be summarized for study participants exposed for given durations of time. For the cumulative duration through Week 48 the following categories for duration will be used:

- >0 weeks
- ≥4 weeks
- ≥8 weeks
- ≥12 weeks
- ≥16 weeks
- ≥20 weeks
- ≥24 weeks
- ≥28 weeks
- ≥32 weeks
- ≥40 weeks
- ≥48 weeks

Definitions for study medication duration and time at risk in days are provided below for each period. Time at risk will be summarized in years. Time at risk in years is calculated by dividing the time at risk in days by 365.25.

Throughout this section, date of last clinical contact for each participant is defined as the maximum of (last visit date including SFU visit, last imputed AE start date, date of study termination or completion, last date of study drug administration).

10.1.1 Exposure during the Initial Treatment Period

Definitions for study medication duration (days) and time at risk (days) during the Initial Treatment Period are provided as follows:

10.1.1.1 Study medication duration (days)

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Period + 14 days.

Note: The use of 14 days assumes a Q2W dosing interval (bimekizumab 320mg Q2W and placebo). For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Initial Treatment Period – Date of first dose in the Initial Period + 28 days).

Note: If date of last dose in the Initial Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Treatment Period + 1.
- For participants who die during the Initial Period, if date of last dose in the Initial Period + 14 days (or date of last bimekizumab dose in the Initial Treatment Period + 28 days in the case of Q4W dosing) extends to a date beyond the date of death, then this calculation reverts to:
 - Date of death – Date of first dose in the Initial Period + 1.

10.1.1.2 Time at risk (days)

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Week 16 visit and continue to the Maintenance Treatment Period:
 - Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Period + 1.
- For participants who discontinue on or prior to the final visit of the Initial Period, use the minimum of the following:
 - Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Treatment Period + 141
 - The total number of days in the Initial Treatment Period (112 days). For AEs that emerged after 112 days but still within the 140 days window, those AEs would be

classified as TEAE, but will be excluded from the output based on the Initial Treatment Period. However, these AEs will be included in the AE summaries for Maintenance Treatment Period.

- Date of last clinical contact – Date of first dose in the Initial Treatment Period + 1.
- For participants who die prior to the final visit of the Initial Treatment Period: Date of death – date of first dose in the Initial Period + 1.

10.1.2 Exposure during the Maintenance Treatment Period

Definitions for study medication duration (days) and time at risk (days) during the Maintenance Treatment Period are provided as follows:

10.1.2.1 Study medication duration (days)

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 14 days.

The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W in the Maintenance Treatment Period, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 28 days).

Note: If date of last dose in the Maintenance Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the final visit date of the Maintenance Treatment Period (not including SFU), then this calculation reverts to:

- Final visit date of the Maintenance Treatment Period (not including SFU) – date of first dose in the Maintenance Treatment Period + 1.
- For participants who die during the Maintenance Treatment Period, then this calculation reverts to:
 - Date of death – Date of first dose in the Maintenance Treatment Period + 1.

10.1.2.2 Time at risk (days)

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study):
 - Date of last visit of the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + 1.
- For participants who die prior to the final visit of the Maintenance Treatment Period:
 - Date of death – Date of first dose in the Maintenance Period + 1.
- For all other participants, use the minimum of the following:
 - Date of last dose in the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + 141 days.

- Date of last clinical contact – Date of first dose in the Maintenance Treatment Period + 1.

Note: This group could include participants who discontinue the Maintenance Treatment Period early, participants who complete the Maintenance Treatment Period as scheduled but choose not to continue into an extension study, or participants who are ongoing in the SFU period at the time of the data snapshot.

10.1.3 Exposure during the Initial and Maintenance Treatment Period

Definitions for study medication duration (days) and time at risk (days) during the Initial and the Maintenance Treatment Period are provided as follows:

10.1.3.1 Study medication duration (days)

Definitions for study medication duration (days) are provided as follows:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Period.

Note: The algorithms for calculating these durations are specified in Section 10.1.1.1 and Section 10.1.2.1.

Note: If date of last dose in the Initial Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Periods - 1.

10.1.3.2 Time at risk (days)

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study):
 - Final visit date – Date of first dose + 1.
- For participants who die prior to the final visit:
 - Date of death – Date of first dose in the + 1.
- For all other participants, use the minimum of the following:
 - Date of last dose – Date of first dose + 141 days.
 - Date of last clinical contact – Date of first dose + 1.

Note: This group could include participants who discontinue early, participants who complete the Maintenance Treatment Period as scheduled but choose not to continue into an extension study, or participants who are ongoing in the SFU period at the time of the data snapshot (in the case of the interim analysis).

10.2 Adverse events

An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of IMP, whether or not considered related to the IMP. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of IMP.

10.2.1 Data considerations

Treatment-emergent AEs are defined as those AEs that have a start date on or following the first dose of study treatment through the final dose of study treatment + 140 days (covering the 20-week SFU period). If it is not possible (due to partial dates) to determine whether or not an AE is treatment-emergent then it will be assumed to be a TEAE.

An AE will be assigned to the Initial Treatment Period if it started between the first administration of IMP on Day 1 up to Week 16. An AE will be assigned to the Maintenance Treatment Period if it started between the Week 16 study drug administration and Week 48.

If an AE occurs on the date of a treatment switch, the event is attributed to the original treatment. The only exception to this is if the AE fulfills any of the criteria specified below:

- Events that fulfill the anaphylaxis criteria for acute events (refer to Appendix A in Section 12.1)
- Hypersensitivity events identified by the SMQ “Hypersensitivity (SMQ)” (see Section 12.1 Appendix A)
- Events with an high level term (HLT) of “Administration site reactions NEC”
- Events with an HLT of “Injection site reactions”

The rules for imputing partial start or stop dates are outlined in Section 4.2.4.

Any TEAEs that occur during the SFU Period will be attributed to the period in which the participant was before initiating the SFU Period.

Duration of AEs will not be calculated if there is missing stop date information.

If the intensity of an AE is unknown, it is considered as severe. If the relationship to study drug is missing, it is considered as related. Note that if the seriousness of an adverse event is unknown, every attempt should be made to resolve this prior to a snapshot for an interim analysis or database lock; in the exceptional case that the seriousness of an adverse event is still missing then no imputation should be applied for this characteristic.

Adverse events will be presented as “number of participants (percentage of participants) [number of events]”. In this style of output, “[number of events]” will include all cases of an AE including repeat occurrences in individual participants, while “number of participants” will count each participant only once.

Subject time at risk represents the time a participant is at risk for having an AE. The definitions for subject time at risk (in days) are outlined in Section 10.1. These definitions will be used for exposure-adjusted AE summaries.

Selected AE summaries will include the exposure-adjusted incident rate (EAIR) with associated 95% CI and the exposure adjusted event rate (EAER).

The EAIR is defined as the number of participants (n) with a specific AE adjusted for the exposure and will be scaled to 100 subject-years:

$$EAIR = 100 \times n / \sum_{i=1}^N (T_{Exp(i)})$$

Where $T_{Exp(i)}$ is the exposure time and N is the number of participants at risk.

If a participant has multiple events, the time of exposure is calculated to the first occurrence of the AE being considered. If a participant has no events, the total time at risk is used.

Exact Poisson 95% confidence intervals for incidence rates are calculated using the relationship between the Poisson and the Chi-square distribution (Ulm, 1990; Fay and Feuer, 1997):

$$LCL = \chi^2_{2n, \alpha/2} / 2$$

$$UCL = \chi^2_{2(n+1), 1-\alpha/2} / 2$$

where n is the number of participants with a specific AE for the incidence rate of interest and is the basis for the number of the degrees of freedom for the chi-square quantile for the upper tail probability χ^2 .

The EAER will be the number of AEs including repeat occurrences in individual participants divided by the total time at risk scaled to 100 patient-years and calculated using:

$$EAER = 100 \times N_{AE} / \sum_{i=1}^N (T_{Risk(i)})$$

where N_{AE} is the total number of AEs, T_{Risk} is the time at risk for each participant, and N is the total number of participants at risk.

No confidence interval will be computed for EAER.

Selected summaries, as specified in Section 10.2.2, will include the risk difference between bimekizumab and placebo. The risk difference is calculated as:

$$RD = IP_{BKZ} - IP_{PBO}$$

where IP_{BKZ} is the incidence proportion for the bimekizumab-treated group and IP_{PBO} is the incidence proportion for the placebo group. Note that incidence proportion simply refers to the percentage of participants within the specified treatment group that experienced a given adverse event.

The standard error for the risk difference is calculated as follows:

$$SE_{RD} = \sqrt{\left(IP_{BKZ} \times \left(\frac{1 - IP_{BKZ}}{n_{BKZ}} \right) \right) + \left(IP_{PBO} \times \left(\frac{1 - IP_{PBO}}{n_{PBO}} \right) \right)}$$

where n_{BKZ} is the number of participants in the bimekizumab-treated group and n_{PBO} is the number of participants in the placebo group.

The corresponding confidence interval for the risk difference is as follows:

$$CI_{RD} = RD \pm Z_{1-\alpha/2} \times (SE_{RD})$$

where $Z_{1-\alpha/2}$ is the Z statistic for the corresponding level of alpha. For the risk difference confidence intervals calculated in this SAP, 1.96 will be used (corresponding to a two-sided alpha of 0.05 and 95% confidence interval). The risk difference and corresponding CI will be displayed as percentage.

10.2.1.1 COVID-19 related considerations

To assess the impact of COVID-19 mass vaccination on TEAEs, a sensitivity analysis will present all TEAEs excluding TEAEs assessed as exclusively related to COVID-19 vaccine by the investigator. TEAEs recorded as related to both study medication and COVID-19 vaccination should not be excluded. A complementary table and listing of TEAEs related to COVID-19 vaccine will be presented.

Another sensitivity analysis will present all TEAEs excluding TEAEs with start date on or up to 5 days after date of COVID-19 vaccine. Note that study participants may receive more than one administration of COVID-19 vaccine. A complementary table and listing of TEAEs with start date on or up to 5 days after date of COVID-19 vaccine will also be presented.

10.2.2 AE summaries

The following summaries will be provided by treatment group for the Initial Treatment Period, and the Initial and Maintenance Treatment Period combined based on the SS, and AMS respectively. In addition, all summaries of TEAEs based on “100 subject years” will include EAIR (with 95% confidence interval) and EAER. For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE. These AEs will be excluded from the outputs based on the Initial Period but included in the AE summaries for Initial and Maintenance Treatment Period.

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Study Participant Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT
- Incidence of TEAEs by Maximum Relationship by SOC, HLT, and PT
- Incidence of Serious TEAEs by Relationship SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of Related Serious TEAE by SOC, HLT, and PT
- Incidence of Severe TEAE per 100 subject years by SOC, HLT, and PT

- Incidence of TEAEs Leading to Death by Relationship by SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of TEAEs by Maximum Severity, SOC, HLT, and PT
- Incidence of TEAEs by decreasing frequency of PT
- Incidence of TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Maximum Relationship SOC, HLT, and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by Relationship SOC and PT
- Incidence of Related TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Related TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of TEAEs – Suspected and Confirmed COVID-19 cases by SOC, HLT and PT
- Incidence of TEAEs Excluding TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval Censored TEAEs by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval TEAEs by SOC, HLT, and PT

Suspected and confirmed COVID-19 cases will be identified with the preferred terms “Corona virus infection” or “Corona virus test positive”.

The following subset of tables will also be presented for the Maintenance Treatment Period using the MS:

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT

The following tables will be presented for the Initial Treatment Period:

- Incidence of TEAEs Above Reporting Threshold of 5% with Risk Differences by SOC and PT
- Incidence of Serious TEAEs and Risk Differences by SOC and PT

The following table will be presented for the combined Initial and Maintenance Treatment Period. This summary will include only AEs that occur while a participant is on bimekizumab.

Any AEs in the Initial Treatment Period that begin while a participant is on placebo will be excluded.

- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status. This will include columns for the following:
 - TEAEs starting before the first ADA b positive result (includes ADA b categories 2 and 5) where TEAEs have occurred before the following events: a) the first positive ADA b result for subjects in category 2 and b) the first post-Baseline boosted ADA b titer result for subjects with titer results and the first post-Baseline positive ADA b result for subjects with positive ADA b at Baseline with no other samples with titer available for subjects in category 5
 - TEAEs starting on the same date or after the first ADA b positive result (includes ADA b Categories 2, 3, 4, 5 and 6) where TEAEs have occurred on or after the following events: a) the first positive ADA b results for subjects in categories 2, 3, 4 and 6, and b) the first post-Baseline boosted ADA b titer result for subjects with titer results and the first post-Baseline positive ADA b result for subjects with positive ADA b at Baseline with no other samples with titer available for subjects in category 5
 - TEAEs for subjects who are ADA b negative at all timepoints (includes ADA b Category 1)

The tables with risk differences will also be accompanied by figures (dot plots) which show the incidence of the adverse events and corresponding 95% risk difference confidence intervals. These will be ordered by descending order of risk difference (bimekizumab vs placebo).

10.2.3 Other Safety topics of interest

The following are AEs considered to be other safety topics of interest that require special statistical analyses. Along with the tables described, there will be a table which displays the risk difference and 95% confidence intervals for each of the topics of interest in the Initial Treatment Period. A corresponding figure (with dot plots) will be prepared.

A by-participant listing of all AEs of safety topics of interest will be presented by type of safety topics of interest.

10.2.3.1 Infections (serious, opportunistic, fungal and TB)

- **Incidence of Serious Infection TEAEs per 100 subject years by SOC, HLT and PT**

Serious infections will be identified based on MedDRA classification (SOC “Infections and infestations”) using the “Any SAE” table. A separate table does not need to be produced to summarize these events.

- **Incidence of Fungal Infection TEAEs per 100 subject years by SOC, HLT and PT**

Fungal infections will be summarized in a stand-alone table. The table will include all TEAEs (serious and non-serious) that code into the High Level Group Term (HLGT) “Fungal infectious disorders”

- **Incidence of Opportunistic Infection TEAEs per 100 subject years by SOC, HLT and PT**

Opportunistic infections (including tuberculosis) will be summarized in a stand-alone table. The table will include all TEAEs identified using UCB-defined search criteria .

The following steps will be followed for identifying and reviewing opportunistic infections:

Identification Process

The steps below outline 2 ways in which opportunistic infections (or potential opportunistic infections) can be identified:

Step 1: Refer to column B of the spreadsheet, which identifies the PTs to be classified as opportunistic infections using either a single 'x' or a double 'xx'.

- All TEAEs that code to a PT flagged with a single 'x' need to also be serious to be considered an opportunistic infection.
- All TEAEs that code to a PT flagged with a double 'xx' are considered to be an opportunistic infection, regardless of seriousness.

Step 2: Refer to column C of the spreadsheet, which identifies the PTs that need to be evaluated on a case-by-case basis by the study physician to determine whether or not it is an opportunistic infection. If column C has a single 'x', then the corresponding preferred term will be flagged for case-by-case review by the study physician.

Review Process

Opportunistic infections for a given study will be reviewed on the following occasions:

- At quarterly Infectious Disease Committee (IDC) Meetings, listings will be produced for each study (see details below) and reviewed by the corresponding study physician ahead of the IDC Meeting. These listings will be posted as part of the broader Safety Signal Detection (SSD) deliverable to a folder named for the given quarter (eg, 2018Q4) on the SharePoint. They will be based on the same data cut as the one used for SSD and will be delivered at the same time as the SSD outputs. The IDC will then agree on the final adjudication for each potential opportunistic infection.
- For each study, a final listing for opportunistic infections (in the format described below) will be produced and agreed upon between the study physician and the IDC prior to finalizing the database.

In each of the circumstances described above, the study programming team will produce an Excel listing that will be provided to the project lead statistician, project lead programmer, and to the study physician (who will subsequently provide it to the IDC). The Excel listing will contain the following columns (using the descriptions below as the column headings in the Excel listing):

- Study ID
- Unique Participant ID
- AE Term (Verbatim)
- AE Preferred Term
- AE System Organ Class
- AE High Level Term

- AE Low Level Term
- Date of Onset
- Outcome of Adverse Event
- Date of Outcome
- TEAE Flag
- Serious Adverse Event?
- Relationship to Study Medication
- Intensity
- Action Taken with IMP
- Opportunistic Infection – Automatic
- Opportunistic Infection – Manual Review
- Flag
- Data Cut Date
- Opportunistic Infection – Final Adjudication

Note the following about the final 5 variables in this listing:

- *Opportunistic Infection – Automatic*: This is flagged as “Y” if the criteria for automatic selection as described in “Step 1” of the identification process are met.
- *Opportunistic Infection – Manual Review*: This is flagged as “Y” if the criteria for case-by-case selection as described in “Step 2” of the identification process are met.
- *Flag* – This has a value of either “NEW” or “OLD”. It is marked as “NEW” if the event is appearing for the first time in that run of the listing. Otherwise, if it has appeared previously, it is marked as “OLD”. Unique records are determined by USUBJID AESPID for purposes of identifying whether an event has been modified from a previous run.
- *Date* – Only for cases where Flag is “NEW”, this field will be populated with the data cut date for that particular run of the listing.
- *Opportunistic Infection – Final Adjudication* – For new events, this is always left blank by the programmers. It will be completed by the study physician/IDC for every event that appears in the listing. For events adjudicated as opportunistic, the field will be populated with a “Y”.

Following each review by the study physician and IDC, the Opportunistic Infection – Final Adjudication column will be completed (as described above), and the spreadsheets for each study will be returned to the study programming team via e-mail (coordinated by the IDC secretary). Then, for subsequent runs of the listing, the study programming teams will incorporate adjudications from previous runs.

10.2.3.2 Malignancies

- **Incidence of Malignant or Unspecified Tumours TEAEs per 100 subject years by SOC, HLT and PT**

These events will be presented in the following tables:

- One table will be based on the criteria Standardized MedDRA Query (SMQ) = “Malignant or unspecified tumours (SMQ)”
- One table will be based on the criteria SMQ = “Malignant tumours (SMQ)”.

SMQ search will include all TEAEs which code to a PT included in the Scope=Narrow group within each SMQ.

Note that the events included in the “Malignancies” table will be a subset of the events included in the “Malignancies (including unspecified)” table. While the “Malignant tumours (SMQ)” is most relevant, “Malignant or unspecified tumours (SMQ)” must be reviewed for potential malignancies.

The output tables will include 2 different overall incidence rows:

- The first overall incidence row will summarize “Any Malignancy” and this row will summarize the incidence of all AEs flagged for inclusion in the table (using the appropriate SMQ depending on the table), regardless of the HLT it codes to.
- The second overall incidence row will summarize “Any Malignancy excluding non melanomic skin cancers HLT” and this row will summarize the incidence of AEs flagged for inclusion in the table (using the appropriate SMQ depending on the table), excluding those which code to an HLT of “skin neoplasms malignant and unspecified (excl melanoma)”.

10.2.3.3 Major adverse cardiac event

- **Incidence of Adjudicated Major Cardiac Event (MACE) TEAEs per 100 subject years by SOC, HLT and PT**

Potential cardiovascular events are adjudicated by the independent Cardiovascular Event Adjudication Committee (CV-CAC) according to the CV-CAC Charter (version 6.0).

Adjudicated events are classified by the CV-CAC to one of the event types as defined in [Table 10–1](#). The classification of an event as a Major Adverse Cardiac Event (MACE) is also determined by the CV-CAC. Events which are classified by the CV-CAC as any of the event types identified in the third column of [Table 10–1](#) will be considered an extended MACE. Note that extended MACE is determined programmatically and includes a broader scope definition of MACE.

MACE as determined by the CV-CAC will be presented in a table and listing. Extended MACE will be presented separately in a table and listing.

Another table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be a MACE and/or an extended MACE.

Additionally, a listing of all events identified for potential review by the CV-CAC will be produced. This listing will indicate whether each event was identified by the CV-CAC Chair for full committee review.

Table 10–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
1	Non-Fatal Myocardial Infarction (MI)	Yes
2	Non-Fatal Stroke: hemorrhagic	Yes
3	Non-Fatal Stroke: ischemic	Yes
4	Non-Fatal Stroke: embolic	Yes
5	Non-Fatal Stroke: undeterminable	Yes
6	Hospitalization or ER for Unstable Angina with urgent revascularization	Yes
7	Hospitalization or ER for Unstable Angina without urgent revascularization	No
8	Hospitalization for Heart Failure	Yes
9	Transient Ischemic Attack (TIA)	No
10	Coronary Revascularization Procedures (e.g. percutaneous coronary intervention, coronary artery bypass grafting)	Yes
11	Urgent Revascularization Procedures (i.e. due to symptoms of brain ischemia or pending infarction)	Yes
12	Arrhythmia (not associated with ischemia)	No
13	Peripheral Arterial Event	No
14	Venous Thromboembolic Event: pulmonary embolism (PE)	No
15	Venous Thromboembolic Event: deep vein thrombosis (DVT)	No
16	Venous Thromboembolic Event: PE and DVT	No
17	Other CV Event	No
18	Death due to Myocardial Infarction (MI)	Yes
19	Death due to Stroke	Yes
20	Sudden Cardiac Death	Yes

Table 10–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
21	Other CV Death (e.g. heart failure, pulmonary embolism, cardiovascular procedure-related)	Yes
22	Cardiovascular: Undetermined Cause of Death (i.e. cause of death unknown)	Yes
23	Non-Cardiovascular Death	No
24	Non-Cardiovascular Event	No
99	Inadequate information to adjudicate	No

CV=Cardiovascular; DVT=Deep Vein Thrombosis; ER=Emergency Room; MACE=Major Adverse Cardiac Event; MI=Myocardial Infarction; PE=Pulmonary Embolism; TIA=Transient Ischemic Attack.

MACE is determined by the adjudication committee and is not identified programmatically based on event type.

10.2.3.4 Neutropenia

• Incidence of Neutropenia TEAEs per 100 subject years by SOC, HLT and PT

This table will be based on the following PTs (regardless of seriousness):

- Autoimmune neutropenia
- Band neutrophil count decreased
- Cyclic neutropenia
- Febrile neutropenia
- Idiopathic neutropenia
- Neutropenia
- Neutropenic infection
- Neutropenic sepsis
- Neutrophil count decreased

10.2.3.5 Suicidal Ideation and Behavior

• Incidence of Suicidal Ideation or Behavior TEAEs per 100 subject years by SOC, HLT and PT

Potential neuropsychiatric events are adjudicated by the independent Neuropsychiatric Adjudication Committee according to the Neuropsychiatric Adjudication Committee (version 8.0). Adjudicated events are classified by the Committee as Suicidal or Non-suicidal. Adjudicated events are also further classified by the Committee to one of the event types as defined in [Table 10–2](#). Suicidal Ideation and Behavior (SIB) is defined as events classified by the Committee as Suicidal.

A table and listing will present SIB events.

Another table and listing will present the adjudicated neuropsychiatric events by type. For each neuropsychiatric event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be Suicidal or Non-Suicidal. For event type suicidal ideation, the listing will also indicate if intent was present and if the suicidal ideation was clinically significant.

Additionally, a listing of all events identified for potential review by the Committee will be produced. This listing will indicate whether each event was identified by the Neuropsychiatric Event Adjudication Committee Chair for full committee review.

Table 10–2: Neuropsychiatric event classifications

Event Type Code	Event Classification	Event Type
1	Suicidal	Suicidal events/completed suicide
2	Suicidal	Suicide attempt
3	Suicidal	Preparatory acts toward imminent suicidal behavior
4	Suicidal/Non-suicidal ^a	Suicidal ideation
7	Non-suicidal	Nonsuicidal Self-injurious behavior
8	Non-suicidal	Nonsuicidal Other
99	Not applicable	Inadequate information to adjudicate

^a Suicidal ideation event types can be classified by the Neuropsychiatric Adjudication Committee as Suicidal or Non-suicidal depending on whether intent to die was present.

10.2.3.6 Inflammatory bowel disease

- **Incidence of Inflammatory Bowel Disease TEAEs per 100 subject years by SOC, HLT and PT**

Selected gastrointestinal events are adjudicated by the independent Inflammatory Bowel Disease (IBD) Adjudication Committee (IBD-CAC) according to the IBD-CAC Charter (version 3.0). Adjudicated events are classified by the IBD-CAC into one of the diagnostic types as defined in [Table 10–3](#). The events will further be classified as definite, probable or possible IBD.

An overview of adjudicated IBD events will be stratified by subjects with or without a previous medical history of IBD. Previous medical history of IBD will be determined using the information recorded on the History of IBD CRF page ("Does subject have a history of IBD?"). This overview table will present events adjudicated by the IBD-CAC as either possible, probable or definite IBD. Definite and probable IBD will also be aggregated and summarized in this table. In addition, this table will summarize each IBD event classification (possible, probable or definite) separately.

Another table and listing will present the adjudicated IBD events by type. For each IBD event type, the individual PTs which fall within each event type will be summarized.

Additionally, a listing of all events identified for potential review by the IBD-CAC will be produced. This listing will indicate whether each event was identified by the IBD-CAC Chair for full committee review.

A further supportive listing will present the individual diagnostic criteria met for each adjudicated IBD event.

Table 10–3: IBD event classifications

Event Type Code	Event Type (Classification and diagnosis)	Classification
1	Possible Inflammatory Bowel Disease – Crohn’s Disease	Possible
2	Probable Inflammatory Bowel Disease – Crohn’s Disease	Probable
3	Definite Inflammatory Bowel Disease – Crohn’s Disease	Definite
4	Possible Inflammatory Bowel Disease – Ulcerative Colitis	Possible
5	Probable Inflammatory Bowel Disease – Ulcerative Colitis	Probable
6	Definite Inflammatory Bowel Disease – Ulcerative Colitis	Definite
7	Possible Inflammatory Bowel Disease – type unclassified	Possible
8	Probable Inflammatory Bowel Disease – type unclassified	Probable
9	Definite Inflammatory Bowel Disease – type unclassified	Definite
10	Symptoms not consistent with Inflammatory Bowel Disease	Not applicable
11	Possible Inflammatory Bowel Disease – Microscopic Colitis	Possible
12	Probable Inflammatory Bowel Disease – Microscopic Colitis	Probable
13	Definite Inflammatory Bowel Disease – Microscopic Colitis	Definite
14	Possible Inflammatory Bowel Disease – no further differentiation possible	Possible
15	Probable Inflammatory Bowel Disease – no further differentiation possible	Probable
16	Definite Inflammatory Bowel Disease – no further differentiation possible	Definite
99	Not enough information to adjudicate	Not applicable

IBD=inflammatory bowel disease.

Note: IBD diagnoses of “microscopic colitis” and “no further differentiation possible” were added in an adjudication charter amendment, accounting for the event type numbering.

10.2.3.7 Hypersensitivity (including anaphylaxis)

- Incidence of Anaphylactic Reaction TEAEs per 100 subject years by SOC, HLT and PT

A separate table will be prepared based on the MedDRA anaphylaxis Algorithm (see Appendix 1) for acute anaphylactic events (reported on the same day as when an injection was administered or 1 day after). An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

A separate table will be prepared to summarize hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. In addition, a separate table will be prepared to summarize serious hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All serious TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

Furthermore, a separate table will be prepared to summarize injection site reactions, identified using the HLTs: “Administration site reactions NEC” and “Injection site reactions”.

10.2.3.8 Hepatic events and PDILI

- **Incidence of hepatic events TEAEs per 100 subject years by SOC, HLT and PT**

A table for hepatic events will be created based on the SMQ of “Drug related hepatic disorders - comprehensive search (SMQ)”. However, these 2 sub-SMQs are to be excluded: “Liver neoplasms, benign (incl cysts and polyps) (SMQ)” and “Liver neoplasms, malignant and unspecified (SMQ)”. For each of the above SMQs, include all TEAEs which code to a PT included in the Scope=Broad and/or Scope=Narrow.

Note that all AEs meeting the above criteria are to be included. It will not be limited to events that the investigator determined to be related to study drug.

Cases of potential Hy’s Law will be reported separately in a liver function test table.

10.3 Clinical laboratory evaluations

Laboratory values, including markedly abnormal laboratory values will be presented descriptively by treatment group for the SS, MS, and AMS.

The markedly abnormal tables and those based on common terminology criteria for AEs (CTCAE) grade will be produced only for selected laboratory variables.

For tables where data are summarized by visit, unscheduled and repeat visits will not be summarized, but these data will be included in listings. For tables where multiple measurements over a period of time are considered (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered). All summaries will be presented in SI units and will be based on observed case values. In the case where laboratory values are below the LLQ, then these will be set to the midpoint between 0 and the lower limit of quantification for the purpose of summarizing the data. The following summaries are required:

- A summary of the absolute and change from Baseline values in each laboratory variable by treatment group and visit

- A summary of the number and percentage of participants experiencing markedly abnormal values at any time while on treatment (assessment on or following the first dose of study treatment through the minimum of period of interest (Week 16) or date of last dose + 140 days) by laboratory variable and treatment group. Two separate tables will show results for the Initial Treatment Period (for the SS) and the Initial and Maintenance Treatment Period (for the AMS).
- A summary of the number and percentage of participants with a given CTCAE grade (0,1,2,3, or 4) based on minimum/maximum post-baseline value by laboratory variable and treatment group. Two separate tables will show results for the Initial Treatment Period (for the SS) and the Initial and Maintenance Treatment Period (for the AMS).
- A shift table of the number and percentage of participants experiencing CTCAE grade 0,1,2,3, or 4 values (as applicable) at Baseline to maximum post-Baseline CTCAE grade, by laboratory variable and treatment group. Two separate tables will show results for the Initial Treatment Period (for the SS) and the Initial and Maintenance Treatment Period (for the AMS).
- A by-participant listing of all laboratory data (including urinalysis) will be provided. This listing will be presented by treatment group and will include: center, participant identifier, age, sex, race, weight, visit, laboratory variable, result (with abnormal values flagged as “L” or “H” accordingly) and unit.

Markedly abnormal values are defined as those with a severity of Grade 3 and above based on the CTCAE criteria Version 4.03. Definitions of markedly abnormal values using the Grade 3 cut points are given in the tables below for age ranges of ≥ 17 years [Table 10–4](#) for markedly abnormal liver function test values, [Table 10–5](#) for markedly abnormal biochemistry values and [Table 10–6](#) for markedly abnormal hematology values). Tables summarizing markedly abnormal values will include a summary (counts and percentages) of markedly abnormal labs observed at any time while on treatment (ie, treatment-emergent markedly abnormal [TEMA]). For this summary, Baseline values and values observed more than 140 days after the last administration of study medication are not considered. The laboratory results classified as Grade 3 or Grade 4 will be summarized and listed separately.

Table 10–4: Definitions of Markedly Abnormal Liver Function Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
ALP	U/L	$>5.0 \times \text{ULN}$	U/L	$>5.0 \times \text{ULN}$	AH
ALT	U/L	$>5.0 \times \text{ULN}$	U/L	$>5.0 \times \text{ULN}$	AH
AST	U/L	$>5.0 \times \text{ULN}$	U/L	$>5.0 \times \text{ULN}$	AH
Total Bilirubin	mg/dL	$>3.0 \times \text{ULN}$	umol/L	$>3.0 \times \text{ULN}$	AH
GGT	U/L	$>5.0 \times \text{ULN}$	U/L	$>5.0 \times \text{ULN}$	AH

Table 10–5: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Creatinine ¹	mg/dL	> 3.0 x Baseline or >3.0 x ULN	umol/L	> 3.0 x Baseline or >3.0 x ULN	AH
Glucose	mg/dL	<40 >250	mmol/L	<1.7 >13.9	AL AH
Calcium	mg/dL	>12.5 <7.0	mmol/L	>3.1 <1.75	AH AL
Magnesium	mg/dL	>3.0 <0.9	mmol/L	>1.23 <0.4	AH AL
Potassium	mmol/L	>6.0 <3.0	mmol/L	>6.0 <3.0	AH AL
Sodium	mmol/L	>155 <130	mmol/L	>155 <130	AH AL
Cholesterol	mg/dL	>400	mmol/L	>10.34	AH

¹ The markedly abnormal definitions for creatinine are based on the logical or, if either criterion is met the creatinine value will be designated as abnormal high.

Table 10–6: Definitions of Markedly Abnormal Hematology Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Hemoglobin	g/dL	<8.0 >4.0 above ULN	g/L	<80 >40 above ULN	AL AH
Lymphocytes Absolute	10 ⁹ /L	<0.5 >20.0	10 ⁹ /L	<0.5 >20.0	AL AH
Neutrophils Absolute	10 ⁹ /L	<1.0	10 ⁹ /L	<1.0	AL
Platelets	10 ⁹ /L	<50	10 ⁹ /L	<50	AL
WBC/Leukocytes	10 ⁹ /L	<2.0 >100	10 ⁹ /L	<2.0 >100	AL AH

Abbreviations: AH=abnormal high; AL=abnormal low; ALP = alkaline phosphatase; ALT= alanine aminotransferase; AST = aspartate aminotransferase; dL = deciliter; GGT: gamma-glutamyltransferase; L = liter; mg = milligram; mmol = millimoles; µg = microgram; ULN = upper limit of normal, WBC=white blood cells.

The table for markedly abnormal liver function tests (LFTs) will contain data beyond the CTCAE Grade 3 thresholds outlined in [Table 10–4](#) above in order to allow for a more thorough review of elevated LFTs. There will be 1 table, which will list the count and percentage of participants meeting the below criteria at any time during the study:

- AST: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- AST or ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- Total Bilirubin: >1.5xULN, >2xULN
- ALP: >1.5xULN

For any participant with at least one markedly abnormal LFT (AST >3xULN, ALT >3xULN, bilirubin >3xULN, or ALP >1.5xULN) the New Ratio (nR) will be calculated as the ratio of either ALT or AST (whichever is higher) to ALP, all expressed as multiples of their ULN as follows:

- $nR = [\text{maximum}(\text{AST/ULN or ALT/ULN})]/(\text{ALP/ULN})$

Any pDILI will be summarized (all criteria must be met at the same assessment):

- (AST or ALT > 3xULN) and Total Bilirubin > 1.5xULN
- (AST or ALT > 3xULN) and Total Bilirubin > 2xULN

In addition, a table will be produced to summarize potential Hy's Law cases. The following definition will be used in that table:

- $[\text{AST} \geq 3\text{xULN or ALT} \geq 3\text{xULN}]$ and Total Bilirubin $\geq 2\text{xULN}$ in the absence of ALP $\geq 2\text{xULN}$

In order to meet the above potential Hy's Law criteria, a participant must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation, if applicable) at the same assessment. For example, a participant who experiences a ≥ 2 x ULN elevation of bilirubin at one visit and a $\geq 3\text{xULN}$ elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's Law criteria.

Potential hepatotoxicity (meeting one of the PDILI or Hy's Law laboratory criteria at least once) will be considered with and without symptoms potentially associated with hepatitis or hypersensitivity according to the investigator (reported on the Symptoms of Hepatitis and Hypersensitivity CRF page).

10.4 Vital signs, physical findings, and other observations related to safety

10.4.1 Vital signs

The following vital signs variables will be summarized: systolic blood pressure (mmHg), diastolic blood pressure (mmHg), body temperature (°C) and heart rate (beats/min). The following summaries will be provided for the SS:

- A summary of the absolute and change from Baseline value for each vital sign variable by treatment group and visit.
- A summary of the number and percentage of participants experiencing at least 1 markedly abnormal value for a vital sign variable as defined in Table 10–7, by treatment group and period (Initial Treatment Period [SS], and Initial and Maintenance Treatment Period [AMS]).

Unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

Table 10–7: Definitions of Markedly Abnormal Blood Pressure Values

Parameter (unit)	Markedly Abnormal Low	Markedly Abnormal High
Systolic blood pressure (mmHg)	<90 and a decrease from Baseline of ≥ 20	>180 and an increase from Baseline of ≥ 20
Diastolic blood pressure (mmHg)	<50 and a decrease from Baseline of ≥ 15	>105 and an increase from Baseline of ≥ 15

A by-participant listing of all vital signs data will be provided. This listing will be presented by treatment group and will include: center, participant identifier, age, sex, race, weight, visit, vital sign variable and result (with abnormal values flagged as “L” or “H” accordingly).

10.4.2 Electrocardiograms

Electrocardiogram data will be analyzed by treatment group and visit for the SS.

A summary of the number and percentage of participants with normal, abnormal ECG results, as determined by the central reader, will be presented for all applicable visits.

The following ECG variables will be summarized (absolute values and change from Baseline) by visit: QT corrected for heart rate using Friderica’s formula (QTcF), RR, PR, QRS and QT.

QTc outliers are defined as QTcF values following dosing that are greater than 450 ms or are increases from Baseline greater than 30 ms. QTcF outliers will be highlighted in the data listing and summarized using the following categories:

- Values >450 ms, >480 ms, >500 ms
- Increase from Baseline of >30 ms, increase from Baseline of >60 ms, increase from Baseline of >90 ms
- Values >450 ms and increases of >30 ms. Values >500 ms and increases of >60 ms

The number and percentage of participant who meet the ECG outlier criteria at any assessment post first dose will be summarized for each period.

Two separate by-participant listing of all 12-lead ECG data will be provided based on interpretation from central reader and from site, respectively.

10.4.3 Other safety endpoints

For by-visit summaries, unscheduled and repeat visits will not be summarized, but these data will be included in listings. By-visit tables should include the SFU visit. Summaries over a period of time (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

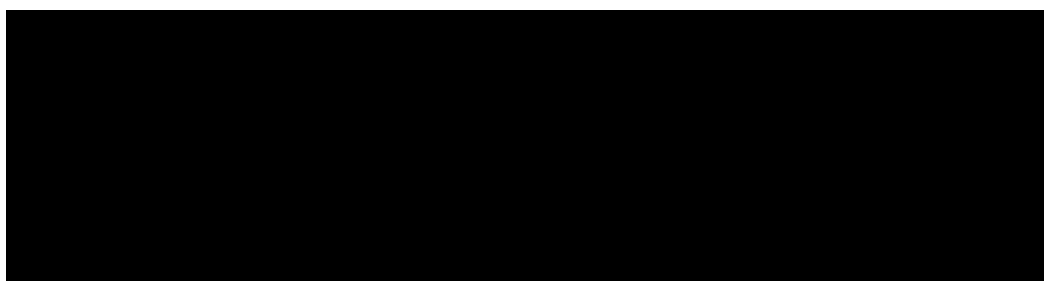
10.4.3.1 Physical examination

Abnormal results of the physical examination together with details of abnormalities: abnormality clinically significant or not, will be listed by participant and visit for SS.

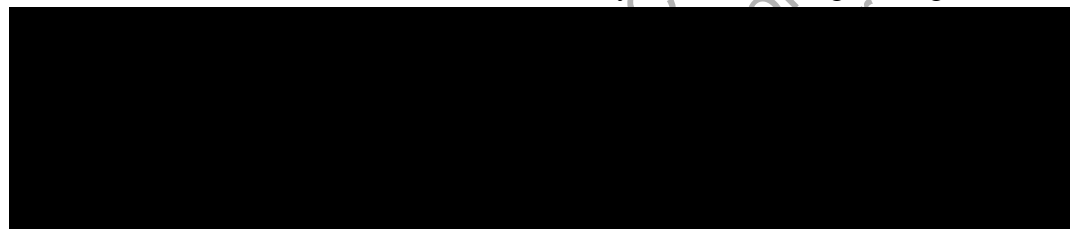
10.4.3.2 Columbia-Suicide Severity Rating Scale (C-SSRS)

The eC-SSRS questionnaire will be self-administered by the study participant and assessed by trained study personnel. This scale will be used to assess SIB that may occur during the study. Results of the eC-SSRS will be summarized using the number of participants and percentage with (i) suicidal ideation, (ii) suicidal behavior, (iii) suicidal ideation or behavior, and (iv) self-injurious behavior without suicidal intent.

Suicidal ideation is defined as an event in any of the following 5 categories:



Suicidal behavior is defined as an event in any of the following 4 categories:



Suicidal behavior or ideation is defined as an event in any of the above 9 categories.

Self-injurious behavior without suicidal intent is defined as an event in the category non-suicidal self-injurious injuries.

The incidence of participants with suicidal ideation, suicidal behavior, suicidal ideation or behavior, and self-injurious behavior will be summarized for the Initial Treatment Period and the combined Initial and Maintenance Treatment Period by treatment group.

A by-participant listing of the eC-SSRS questionnaire data will be provided by treatment group.

10.4.3.3 Assessment and management of TB and TB risk factors

A summary of the number and percentage of participants with negative, positive, and indeterminate IGRA (Interferon-Gamma Release Assay) results at Screening and Week 44 will be presented.

A by-participant listing of the “Evaluation of signs and symptoms of tuberculosis” questionnaire data and IGRA results will be provided by treatment group.

By-participant listing of the result of chest x-ray for tuberculosis will be provided by treatment group.

10.4.3.4 Pregnancy testing

Pregnancy testing will consist of serum testing at the Screening. The pregnancy test will be urine at all other visits.

A by-participant listing of the pregnancy test data will be provided by treatment group.

10.4.3.5 Patient Health Questionnaire (PHQ)-9 scores

The PHQ-9 is a multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression. The PHQ-9 scores for depression range from 0 to 27 with higher scores indicating worse state. A score of 5 to 9 is considered to be minimal symptoms of depression. If any of the 9 questions are missing, then the score is treated as missing. A score of 10 to 14 is considered minor depression, dysthymia, or mild major depression. A score of 15 to 19 is considered to indicate moderately severe major depression, and a score ≥ 20 is considered to be severe major depression.

Change from Baseline in PHQ-9 is derived as post-Baseline score minus Baseline score.

A summary of the absolute and change from Baseline value will be presented by treatment group and visit.

The percentage of study participants with scores below 5, between 5 and 9, between 10 and 14, between 15 and 19, and greater than or equal to 20 in PHQ-9 will be summarized as a shift from Baseline by visit and treatment group based on observed values.

The percentage of study participants with scores ≥ 15 at any post-Baseline visit and the number and percentage of study participants with scores ≥ 20 at any post-Baseline visit will be summarized by treatment group based on observed values. This summary will also include the percentage of study participants with increase from baseline ≥ 5 at any post-Baseline visit.

The number and percentage of participants that complete the PHQ-9 will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit.

11 REFERENCES

Common Terminology Criteria for Adverse Events (CTCAE); Version 4.0 June 2010.U.S. Department of Health and Human Services

Dauden E, Griffiths CE, Ortonne JP, Kragballe K, Molta CT, Robertson D, et al. Improvements in patient reported outcomes in moderate to severe psoriasis patients receiving continuous or paused etanercept treatment over 54 weeks: the CRYSTEL study. *J Eur Acad Dermatol Venereol.* 2009;23(12):1374-82.

Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997; 35(11):1095-108.

Esmann S, Dufour DN, Jemec GB. Questionnaire based diagnosis of hidradenitis suppurativa: specificity, sensitivity and positive predictive value of specific diagnostic questions. *Br J Dermatol.* 2010;163(1):102-6.

Guidance for Industry: E9 Statistical Principles for Clinical Trials. US Department of Health and Human Services, Food and Drug Administration. ICH 1998.

Hurley H. Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In: *Dermatologic surgery.* Roenigh R, Roenigh H, editors. Marcel Dekker. 1989:729-39.

Kimball AB, Kerdel F, Adams D, Mrowietz U, Gelfand JM, Gniadecki R. Adalimumab for the treatment of moderate to severe Hidradenitis suppurativa: a parallel randomized trial. *Ann Intern Med.* 2012;157(12):846-55.

Kimball AB, Jemec GB, Yang M, Kageliery A, Signorovitch JE, Okun MM, et al. Assessing the validity, responsiveness and meaningfulness of the Hidradenitis Suppurativa Clinical Response (HiSCR) as the clinical endpoint for hidradenitis suppurativa treatment. *Br J Dermatol.* 2014;171(6):1434-42.

Kimball AB, Okun MM, Williams DA, Gottlieb AB, Papp KA, Zouboulis CC, et al. Two Phase 3 Trials of Adalimumab for Hidradenitis Suppurativa. *N Engl J Med.* 2016a;375(5):422-34.

Kimball AB, Sobell JM, Zouboulis CC, Gu Y, Williams DA, Sundaram M, et al. HiSCR (Hidradenitis Suppurativa Clinical Response): a novel clinical endpoint to evaluate therapeutic outcomes in patients with hidradenitis suppurativa from the placebo controlled portion of a phase 2 adalimumab study. *J Eur Acad Dermatol Venereol.* 2016b;30(6):989-94.

Mallinckrodt CH, Lin Q, Lipkovich I, and Molenberghs G. A structured approach to choosing estimands and estimators in longitudinal clinical trials. *Pharmaceutical Statistics* 2012; 11:456-61.

O'Kelly, Michael, and Bohdana Ratitch. *Clinical trials with missing data: A guide for practitioners.* John Wiley & Sons, 2014

Ratitch, B., Lipkovich, I., O'Kelly, M. Combining Analysis Results from Mutiply Imputed Categorical Data, PharmaSUG, 2013, SP03

Rubin D.B. 1987. *Multiple Imputation for Nonresponse in Surveys.* New York: John Wiley and Sons.

Sartorius K, Emtestam L, Jemec GB, Lapins J. Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *Br J Dermatol*. 2009;161(4):831-9.

Sun H, Snyder E, Koch GG. Statistical planning in confirmatory clinical trials with multiple treatment groups, multiple visits, and multiple endpoints. *J Biopharmaceut Stat*. 2018;28(1):189-211.

UpToDate. 2011. Modified Hidradenitis Suppurativa Score (HSS). [ONLINE] Available at: <http://cursoenarm.net/UPTODATE/contents/mobipreview.htm?36/33/37400>. [Accessed 27 November 2017].

Zouboulis CC, Kyrgidis A, Jemec GBE, Bechara FG, Giamarellos Bourboulis EJ, Ingram JR, et al. Development and validation of IHS4, a novel dynamic scoring system to assess hidradenitis suppurativa/acne inversa severity. *British Journal of Dermatology*. 2017 June 21.

12 APPENDICES

12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The SMQ *Anaphylactic reaction* consists of three parts:

- A **narrow search** containing PTs that represent core anaphylactic reaction terms (Category A – core anaphylactic reaction terms)

Anaphylactic reaction
Anaphylactic shock
Anaphylactic transfusion reaction
Anaphylactoid reaction
Anaphylactoid shock
Circulatory collapse
Dialysis membrane reaction
Kounis syndrome
Shock
Shock symptom
Type I hypersensitivity

- A **broad search** that contains additional terms that are added to those included in the narrow search. These additional terms are signs and symptoms possibly indicative of anaphylactic reaction and categorized in B, C or D

▪ Category B (Upper Airway/Respiratory Terms)

Acute respiratory failure	Nasal obstruction
Asthma	Oedema mouth
Bronchial oedema	Oropharyngeal spasm
Bronchospasm	Oropharyngeal swelling
Cardio-respiratory distress	Respiratory arrest
Chest discomfort	Respiratory distress
Choking	Respiratory failure
Choking sensation	Reversible airways obstruction
Circumoral oedema	Sensation of foreign body
Cough	Sneezing
Cyanosis	Stridor
Dyspnoea	Swollen tongue
Hyperventilation	Tachypnoea
Irregular breathing	Throat tightness
Laryngeal dyspnoea	Tongue oedema

Laryngeal oedema	Tracheal obstruction
Laryngospasm	Tracheal oedema
Laryngotracheal oedema	Upper airway obstruction
Mouth swelling	Wheezing

▪ **Category C (Angioedema/Urticaria/Pruritus/Flush terms)**

Allergic oedema	Oedema
Angioedema	Periorbital oedema
Erythema	Pruritus
Eye oedema	Pruritus allergic
Eye pruritus	Pruritus generalised
Eye swelling	Rash
Eyelid oedema	Rash erythematous
Face oedema	Rash generalised
Flushing	Rash pruritic
Generalised erythema	Skin swelling
Injection site urticaria	Swelling
Lip oedema	Swelling face
Lip swelling	Urticaria
Nodular rash	Urticaria papular
Ocular hyperaemia	

▪ **Category D (Cardiovascular/Hypotension terms)**

Blood pressure decreased
Blood pressure diastolic decreased
Blood pressure systolic decreased
Cardiac arrest
Cardio-respiratory arrest
Cardiovascular insufficiency
Diastolic hypotension
Hypotension

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after, and for scenarios where two events must have been reported, both events must have occurred within one day of each other (as anaphylaxis is an acute event, imputed dates should not be used in the algorithmic approach):

- A narrow term or a term from Category A;
- A term from Category B - (Upper Airway/Respiratory) AND a term from Category C - (Angioedema/Urticaria/Pruritus/Flush);
- A term from Category D - (Cardiovascular/Hypotension) AND [a term from Category B - (Upper Airway/Respiratory) OR a term from Category C - (Angioedema/Urticaria/Pruritus/Flush)]

12.2 Appendix B: Definition of CTCAE grades

Table 12–1: Definition of CTCAE grades by biochemistry parameters						
Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine ¹	High	umol/L	>1-1.5x Baseline or >ULN-1.5 x ULN	>1.5-3.0x Baseline or >1.5 – 3.0 x ULN	>3.0x Baseline or >3.0 – 6.0 x ULN	>6.0 x ULN
Sodium	Low	mmol/L	130-<LLN	N/A	120-<130	<120
Sodium	High	mmol/L	>ULN-150	>150-155	>155-160	>160
Potassium ²	Low	mmol/L	3.0-<LLN	3.0-<LLN	2.5-<3.0	<2.5
Potassium	High	mmol/L	>ULN-5.5	>5.5-6.0	>6.0-7.0	>7.0
Calcium	Low	mmol/L	2.0-<LLN	1.75-<2.0	1.5-<1.75	<1.5
Calcium	High	mmol/L	>ULN-2.9	>2.9-3.1	>3.1-3.4	>3.4
Magnesium	Low	mmol/L	0.5-<LLN	0.4-<0.5	0.3-<0.4	<0.3
Magnesium	High	mmol/L	>ULN-1.23	N/A	>1.23-3.30	>3.30
Cholesterol	High	mmol/L	>ULN-7.75	>7.75-10.34	>10.34- 12.82	>12.82

1 The CTCAE Grade definitions for creatinine are based on the logical or the highest applicable CTCAE grade should be assigned to a creatinine value.

2 The decreased potassium criterion of 3.0-<LLN is specified for both CTCAE Grade 1 and Grade 2; values meeting this criterion will be counted as Grade 2.

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin	Low	g/L	100-<LLN	80-<100	<80	N/A
Hemoglobin ¹	High	g/L	>0-20 above ULN or >0-20 above Baseline if Baseline is above ULN	>20-40 above ULN or >20-40 above Baseline if Baseline is above ULN	>40 above ULN or >40 above Baseline if Baseline is above ULN	N/A
Platelets	Low	10 ⁹ /L	75-<LLN	50-<75	25-<50	<25

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
WBC	Low	$10^9/L$	$3 < LLN$	$2 < 3$	$1 < 2$	< 1
WBC	High	$10^9/L$	N/A	N/A	> 100	N/A
Lymphocytes	Low	$10^9/L$	$0.8 < LLN$	$0.5 < 0.8$	$0.2 < 0.5$	< 0.2
Lymphocytes	High	$10^9/L$	N/A	$> 4-20$	> 20	N/A
Neutrophils	Low	$10^9/L$	$1.5 < LLN$	$1.0 < 1.5$	$0.5 < 1.0$	< 0.5

LLN=lower limit of normal; N/A=not applicable; ULN=upper limit of normal, WBC=white blood cells

1 The CTCAE Grade definitions to be applied are dependent on the Baseline hemoglobin value. If the baseline value is $> ULN$ then the criteria relative to Baseline is applicable, otherwise the criteria relative to ULN is applicable.

Note that participants who meet the decreased potassium criterion of $3.0 < LLN$, which is specified as the decreased potassium lab criterion for both CTCAE Grade 1 and Grade 2, will be counted as Grade 2.

13 AMENDMENT TO THE STATISTICAL ANALYSIS PLAN

13.1 Amendment 1

Rationale for the amendment

The main purposes of this amendment were:

- General update to analyses to align with protocol amendment 3.
- Procedural clarifications from discussions and feedback provided at meetings
- Update to align with the bimekizumab program standards and safety topics of interest

Modifications and changes

Global Changes

Typos and formatting were updated throughout the document.

Global changes:

The following changes were made throughout the SAP:

- References to Section 3.9 were added through the efficacy sections to clarify definition and handling of intercurrent events
- PRO endpoint terminology was updated to match protocol (eg, Worst Pain score in HSSDD instead of Worst Pain in HSSDD)

Specific changes

In addition to the global changes, the following specific changes have been made (formats as missing spaces or redundant spaces are not listed, typos):

Change #1

The following abbreviations have been added:

CFS	COVID-19 Free Set
COVID-19	coronavirus disease 2019
NI	Negative Immunodepletion
nR	New Ratio
NS	Negative Screen
pDILI	potential drug induced liver injury
PI	Positive Immunodepletion
PS	Positive Screen

Change #2

Section 1 Introduction

The protocols were updated:

The SAP is based on the Protocol Amendment ~~3 2~~, **3 February 2021** ~~16 December 2019~~.

Change #3

Section 2.2 Study endpoints

The following text was added:

The endpoints based on HS Symptom Daily Diary (HSSDD) and Hidradenitis Suppurativa Symptom Questionnaire (HSSQ) pain responses are based on the current definitions, which are continuous. It is anticipated that a responder (binary) endpoint will be defined for the HSSDD and HSSQ pain items as well as other symptom items prior to database lock and unblinding, based on separate ongoing, blinded, psychometric analyses aiming to determine threshold for within-patient clinically meaningful improvement.

The below HSSDD and HSSQ pain response endpoints and analyses will be adjusted accordingly in a future SAP amendment.

Change #4

Section 2.2.1.2 Secondary Efficacy endpoints

The secondary endpoints were updated:

The secondary efficacy endpoints are defined as:

- HiSCR₇₅ response (defined as at least a 75% reduction from Baseline in the total AN count with no increase from Baseline in abscess or draining tunnel count) at Week 16
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16
- **Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16**
- Absolute change from Baseline in Skin Pain score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HS Symptom Daily Diary (HSSDD)
- ~~Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16~~

Change #5

Section 2.2.1.3 Other efficacy endpoints

The following text was added:

- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) **and Total score**

Change #6

Section 2.2.3.2 Other safety endpoints

The other safety topics of interest were updated:

- Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (**including anaphylaxis**), suicidal ideation and behavior, ~~depression~~, major **adverse** cardiovascular events, **hepatic events and potential drug-induced liver injury (PDILI)**, ~~function test changes/enzyme elevations~~, malignancies, and inflammatory bowel disease.

Change #7

Section 2.3.1 Study description

The following sentence was deleted:

Enrollment of study participants currently using antibiotics will be capped at 30% of overall enrollment.

Change #8

Section 3.1 General presentation of summaries and analyses

The following text was added:

For PRO continuous variables, descriptive statistics will also include variable score, absolute and percentage changes from baseline, Q1 and Q3, 10th, and 90th percentiles.

If no participants have data at a given time point, then only n=0 will be presented. The other descriptive statistics will be left blank. If n < 3 then the n, minimum, and maximum only will be presented. The other descriptive statistics will be left blank. If n = 3 n, mean, median, minimum and maximum will be presented only. The other descriptive statistics will be left blank.

For categorical variables, the number and percentage of study participants in each category will be presented. Unless otherwise noted, the denominator for percentages will be based on the number of study participants included in the respective analysis set. Study participants with missing data will be accounted for using the following approaches:

- For summaries of demographics and Baseline characteristics: percentages will be based on all study participants in the analysis set and a “Missing” category (corresponding to study participants with missing data for the variable being summarized) will be included as the last row in the list of categories being summarized.
- For summaries of efficacy and safety endpoints, unless otherwise specified: percentages will be based only on those study participants with observed data for the variable being summarized. As the denominator may be different from the number of study participants in the analysis set being considered, the denominator will be displayed in the table. The general format for displaying this will be “n/Nsub (%)”.

Percentages will be presented to 1 decimal place. If the percentage is 100%, a decimal will not be presented. If the count is 0, the percentage will not be presented. Typically, the % sign will be presented in the column header, but not with each individual value.

For the purpose of the tabulations the lower and upper confidence limits for the percentages will be truncated at 0 and 100% respectively. **Confidence intervals (CIs) for the response rates in efficacy summaries based on nonresponder imputation (NRI) will be computed using the Wilson approximation.**

Change #9

Section 3.3 Definition of Baseline

The following text was added:

For randomized participants for whom no start date of treatment is available, the Baseline value will be considered as the last available value on or before the randomization date.

Change #10

Section 3.4 Protocol deviations

The following text was added:

Deviations related to the Coronavirus Disease 2019 (COVID-19) global pandemic are unavoidable deviations from the protocol due to confirmed COVID-19 infection, suspected COVID-19 infection, general circumstances around COVID-19 without infection or any other deviation from the protocol due to COVID-19. COVID-19 protocol deviations will also be reviewed separately as part of the ongoing data cleaning process.

Change #11

The following section was added:

COVID-19 Free Set (Section 3.5.9)

The COVID-19 Free Set (CFS) will consist of all study participants randomized into the study and who have no COVID-19 impact through Week 16. This analysis set will be used for sensitivity analysis of the primary efficacy endpoint.

Change #12

Section 3.8 Coding dictionaries

The following sentence was updated:

Medications will be coded according to the World Health Organization Drug Dictionary (WHODD) version **MAR2021 B3 or later**. Medical procedures will not be coded.

Change #13

The following section was added:

Definition of an intercurrent event (Section 3.9)

Handling of intercurrent events is one of the key elements for the analysis of efficacy endpoints.

An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy (See Section 8.2.2).

Receipt of systemic antibiotic rescue medication is defined as initiating any systemic antibiotic on or after Baseline for any reason (including in response to an AE). The only exception to this rule is if a participant randomized to the antibiotic stratum on a tetracycline antibiotic interrupts their stable dose of tetracycline antibiotic during the study and subsequently restarts the same tetracycline antibiotic as confirmed using the coded preferred term. The restarted dose and frequency of the antibiotic must be the same or lower than the regimen prior to the interruption.

The dates of an intercurrent event are as follows:

- For receipt of systemic antibiotic rescue medication: start date of the antibiotic
- For discontinuation of study treatment due to an AE or lack of efficacy: Last study treatment date + 17 days. Note: study treatment discontinuation includes study discontinuation.

The choice of 17 days is intended to capture the interval between dosing and lesion assessments (14 days), as well as the visit window (3 days).

An additional sensitivity analysis will be conducted where missing data due to COVID-19 will be considered an intercurrent event and will be imputed as a nonresponse at that particular visit. This will be identified when there are missing data at a visit that has been impacted by COVID-19 according the COVID-19 impact CRF page. The date of this intercurrent event will be the date of the impacted visit.

Change #14

Section 3.10 Changes to protocol-defined analyses

The following text was added:

The HiSQOL endpoint was clarified to show that there are only 3 domains: symptoms, psychosocial, activities and adaptations **and to add total score**.

Also, the following endpoint was added to the list of protocol endpoints not included in the analysis:

- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48

Change #15

Section 4.1 Adjustments for covariates

The following text was added:

If a participant is stratified in the incorrect stratum (ie, the stratum recorded in the Interactive voice or web Response System differs from the actual stratum the participant belongs to), the actual stratum will be used for the analysis.

Change #16

Section 4.2.1.4 Missing Data Overview and Summary

The following text was added:

In summary, the approaches listed below will be used in this study for handling missing data for efficacy endpoints as appropriate:

- **NRI: Participants** who have missing data at the timepoint of interest are treated as though they did not respond to the treatment. **This approach is also referred to as Composite Estimand (NRI).**
- Multiple Imputation (MI) – MCMC / Monotone Regression: Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using monotone regression.

- **MI-MCMC / Reference-based imputation:** Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using an imputation model based on placebo (reference) data.
- **LOCF:** Post-Baseline missing data are imputed by carrying forward the last available observation (including Baseline).
- **Tipping point analyses:** Assumptions will be made about average outcomes among the subsets of **participants** who prematurely discontinued study treatment. Various “delta adjustments” will be made to the assumed responses among missing data in each treatment arm with varying degrees of plausibility in order to identify assumptions about the missing data under which the conclusions change (O’Kelly, 2014). Then, the plausibility of such assumptions is discussed.
- **Observed case (OC):** Missing data are not imputed. Only **participants** with available data who have not discontinued study treatment at the given timepoint are considered. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing. ~~For OC summaries, intercurrent events are not handled differently than other missing data.~~
- **Treatment policy strategy:** All available data observed at the time point of interest will be considered, regardless of the occurrence of intercurrent events. This means the analysis includes on- and off-treatment values collected after study participants prematurely discontinued study treatment but agreed to remain on the study and continued to attend visits and provide assessments at those visits as well as values from study participants who received rescue antibiotic medication. Those observed values will be analyzed according to the study participant’s randomized treatment. Study participants for whom efficacy data cannot be obtained at the week of interest, despite attempts to retain them in the study, will have their data imputed using MI – MCMC / monotone.

Table 4-1 was updated:

Table 4–1: Missing data handling approach by endpoint priority and type								
Endpoint Priority	Endpoint Type	Composite Estimand (NRI)	Modified Composite Estimand (MI)	MI (MCMC/ Reference-based)	Tipping Point	Treatment Policy	Hypothetical Estimand	OC
Primary	Responder	S ^a	P	S ^a	S	S ^a		S
Secondary included in the statistical testing procedure	Responder	S ^a	P					S
	Continuous						P	S
	Binary	X	X					X

Secondary not included in statistical testing procedure	Continuous						X	X
Other	Responder	X^d	X					X^d
	Continuous						X	X^b
	Ordinal						X^c	X^c

B=Backup method, LOCF=Last observation carried forward, MCMC=Markov Chain Monte Carlo, MI=multiple imputation, NRI=Nonresponder imputation, OC=Observed case, P=Primary method, S=Sensitivity method, **X=Method to be used (no priority designated).**

Note: Composite estimand (NRI) refers Backup method is only applicable when the primary method is unable to the approach in which data preceded by the intercurrent event of study treatment discontinuation converge due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are also imputed as nonresponse.

Note: Modified Composite Estimand (MI) refers to the approach in which data preceded by the intercurrent event of study treatment discontinuation due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are imputed via a multiple imputation model.

Note: Hypothetical Estimand (MI) refers to the approach where outcomes for study participants without an intercurrent event of study treatment discontinuation are as observed, and outcomes for study participants challenges with the intercurrent event are imputed via a multiple imputation model.

^a Imputation method is applied on continuous data, and responder endpoint is derived from the continuous endpoint based on complete data set where applicable.

^b Required only for by-visit summaries of variables whose value at Week 16 is part of the hierarchical testing procedure.

^c For variables with multiple categories, data will be summarized as observed with an additional missing row to capture missing data at a given visit.

^e Participants with intercurrent events are imputed as nonresponders for all subsequent timepoints before the imputation method is applied for all other missing data.

^d NRI/OC sensitivity analysis will be performed only for HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀ summaries.

^e The treatment policy estimand will use the same MI-MCMC/Monotone Regression defined for the primary analysis, with the exception that participants with intercurrent events will not be treated as nonresponders for all subsequent timepoints before the imputation method is applied for missing data.

Change #17

Section 4.2.2 Missing data algorithms for efficacy analyses

The section has been moved from Section 4.2.1.5 and split into Section 4.2.2.1, Section 4.2.2.2, and Section 4.2.2.3.

Change #18

Section 4.2.2.1 MI – MCMC/Monotone Regression

This section was updated to the following text:

In many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent

with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied, as follows:

Binary endpoint

For a binary endpoint (eg, HiSCR₅₀), the procedure is as follows:

6. Create a data set, one for each treatment group of participants with observed values and those needing estimation by multiple imputation. For the imputation step, a distinction is made between non-monotone missing values (ie, intermittent missing values between completed assessments) and monotone missing values (ie, where all participants data are missing after a given time point).
- c. For the intermittent missing values, the missing values in each data set will be filled in using the MCMC method with multiple chain, monotone missing data imputing pattern, and non-informative prior for all parameters. Unless specified differently, the first 200 iterations will not be used (the “Burn-in” option). A total of 100 sets of imputations will be performed. The seed used for these imputations will be 762 and all other multiple imputation procedures described in this SAP will use this same seed as well. The resulting 100 imputed data sets will have a monotone missing pattern and will be imputed using a method for monotone missingness. **Note** Create a data set, one for each treatment group (note that a separate imputation procedure must be invoked in SAS for each treatment group as the seed cannot be set for by groups beyond the first when using a by statement), of participants with observed values and those needing estimation by multiple imputation. The intermittent missing lesion counts in PROC MI.
- d. each data set (ie, missing values for a given subject that has available data before and after the missing timepoint) will be filled in using the MCMC method, with a total of 100 sets of imputations being performed. The seed used for these imputations will be 762 (note that all other multiple imputation procedures described in this SAP related to MCMC/Monotone regression analyses will use this same seed as well). For monotone missing data (ie, where all participant data is missing after a given timepoint), monotone regression will then be used to impute missing data. **A separate regression model is estimated for each variable with missing values (ie, measurement at each time point). Based on the resulting model, a new regression model is then drawn and is used to impute the missing values for the variable. Since the data set has a monotone missing data pattern, the process is repeated sequentially for variables with missing values. This procedure will be based on the 100 imputed datasets generated from sets of imputations already created using the MCMC procedure and method such that there will be performed by imputation. The SAS® PROC MI procedure will be used for the imputation.**
 1. ~~data sets in total.~~ In both cases, Hurley Stage at Baseline, Baseline antibiotic use, and **value of the variable of interest lesion count** at Baseline and at each post-Baseline visit (prior to the time point of interest in chronological order, see notes below about visits to include for different analysis sets) will be included in the imputation model. **The post-Baseline values will need to be specified in chronological order in the imputation model so that SAS® PROC MI imputes variables from left to right (eg, the Week 2 value will be first imputed using regression based on the Baseline value, and then Week 4 value will be imputed using regression based on Baseline and Week 2**

values, etc). Note that lesion count at earlier visits will also be used as predictors for the model of lesion count at later visits. The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model **based on the MCMC method** will only allow **joint multivariate normal numeric variables**. Therefore, Hurley Stage at Baseline and Baseline antibiotic use will be re-coded as indicator variables. For Baseline antibiotic use, this will simply be 0 for Baseline antibiotic non-users and 1 for Baseline antibiotic users. For Hurley Stage at Baseline, this will be 0 for Hurley Stage II participants and 1 for Hurley Stage III participants. In order to achieve model convergence, Baseline antibiotic use may be dropped from the model. If convergence is still not obtained, then Hurley Stage at Baseline may also be dropped from the model. **Additionally, if a variable is dropped in order to allow convergence for one model in a study, that variable does not have to be dropped from other models in the study if the model converges without dropping the variable. In other words, model convergence should be evaluated for each efficacy variable independently.**

Note: The imputation of each lesion type (inflammatory nodule, abscess, draining tunnel, etc) will be performed separately. The 100 data sets obtained for each type will be merged by imputation number and subject number.

- For each complete imputed data set, the dichotomous responder variable (eg, HiSCR 0 or 1) ~~based on the imputed %ΔAN and draining tunnel (fistula/sinus tract) count~~ will be computed. Each complete imputed data set will then be analyzed based on the logistic regression model.

Note: For derivation of HiSCR response, the AN, inflammatory nodule, abscess, and draining tunnel (fistula/sinus tract) counts at Week 16 in the imputed data sets will be compared directly to the observed Baseline counts to determine response. If values outside of the pre-defined range of values for lesion count (<0) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed draining tunnel (fistula/sinus tract) count of **-1** would be changed to 0 before deriving the HiSCR responder variable. Additional ranges for values for secondary and other endpoints are defined in Table 4-2.

Note: Standard rounding rules will also be applied to the imputed values of endpoints that can only take integer values (eg, abscess count). For example, if a study participant has an abscess count imputed as 2.4, this imputed value would be rounded down to 2. This rounding step is performed after the multiple imputation but before deriving the responder variable.

Table 4-2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
Lesion count ^a	0	--	Yes
DLQI total score	0	30	Yes
hs-CRP	LLOQ/2	--	No
HSSDD item score	0	10	Yes

Table 4–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
HSSQ item score	0	10	Yes
HHS4	0	—	No
HiSQOL total score	0	68	Yes
HiSQOL symptom status score	0	16	Yes
HiSQOL psychosocial impact score	0	20	Yes
HiSQOL impact on physical activities score	0	32	Yes
EQ-5D-3L VAS	0	100	Yes
WPAI dimension scores	0	100	No for variables: “Percent work time missed due to problem” and “Percent overall work impairment due to problem”. Yes for variables: “Percent impairment while working due to problem” and “Percent activity impairment due to problem”. These two variables can only take values that are multiples of 10.
PHQ-9	0	27	Yes

^a Lesion counts will be imputed separately for each lesion type (abscesses, draining tunnels [fistulas/sinus tracts], inflammatory nodules, non-draining tunnels [fistulas/sinus tracts], non-inflammatory nodules, HS scars). **The imputed lesion counts will be used to derive the endpoints that are dependent on the lesion count data (eg, HiSCR₅₀).**

7. **Estimates of the adjusted responder rate for each treatment group and the associated SE are obtained** The Week 16 results **from the specified statistical analysis** (logistic regression model per Section 8.2.2) of each of the 100 imputed data sets. **These estimates will be combined for overall inference using Rubin’s rules, which account for the uncertainty associated with the imputed values (Rubin, 1987), and the combined estimates and SEs will be used to construct 95% CIs using the logit scale. This will be done using SAS PROC MIANALYZE. The combined estimates and 95% CIs on the logit scale will be back-transformed using the inverse logit link function to obtain the adjusted responder rates (%) and associated 95% CIs. This will be done using SAS PROC MIANALYZE.**

Note: The (unadjusted) proportion of responders will be calculated at each time point by treatment group from the imputed datasets using SAS PROC FREQ. These results will also be combined into an overall inference using SAS PROC MIANALYZE.

Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. However, in the event that there are computational challenges with the imputation model (eg, due to a standard deviation of

0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

As the estimates of the odds ratios from the logistic regression models in Step 3 follow a log-normal distribution, a log transformation is needed to normalize these 100 odds ratio estimates. That is because the procedures for combining results from multiple imputed datasets assume that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (**Step 3**). **Additionally, the SE for the odds ratios are transformed as follows:** the use of PROC MIANALYZE in step 3). Appropriate transformations to the standard errors and p-values will also be made in order to get the correct confidence intervals. For the logistic regression using the p-value for the general association the Wilson-Hilferty transformation will be used (Ratitch, 2013).

$$SE = \frac{\log(UCL) - \log(LCL)}{2Z_{\alpha/2}}$$

Where UCL and LCL are the upper and lower confidence limit, respectively, for the CI of the odds ratio from the logistic regression model, and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). The estimates of the log odds ratio for Bimekizumab relative to placebo and the corresponding upper and lower CLs will be provided. The odds ratio will be then estimated by exponentiating the estimate of the log odds ratio. The odds ratio and the confidence limits of the odds ratio will be estimated as follows:

$$OR = \exp(\text{Log odds ratio estimate})$$

$$LCL = OR * \exp(-SE * Z_{\alpha/2})$$

$$UCL = OR * \exp(SE * Z_{\alpha/2})$$

Where OR is the back-transformed estimate of the odds ratio just described, SE is the SE of the log odds ratio and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). These calculations will be done such that odds ratios and corresponding CIs are calculated for the odds ratio of bimekizumab vs. placebo.

Note: . If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose, with a corresponding $Z_{\alpha/2}$ of 1.96.

In addition to calculating the odds ratio, associated CIs, and p-values for the pairwise comparisons of bimekizumab and placebo, the estimated proportion of responders (ie, estimated responder rate) and the difference in the proportion of responders between each bimekizumab treatment group and placebo will be estimated, and 2-sided 95% CIs will be

created for each difference. The creation of the estimates of the differences will be completed for each bimekizumab treatment group using the process detailed below:

8. Use the logistic regression model to calculate:

Least squares mean estimates of the log odds of bimekizumab (G_B) and placebo (G_P), as well as their corresponding standard errors (S_B and S_P , respectively).

Standard error of the least squares mean estimate of the log odds ratio (S_R)

9. Compute estimates for predicted proportions using the following transformations:

$$P_B = \exp(G_B) / (1 + \exp(G_B))$$

$$P_P = \exp(G_P) / (1 + \exp(G_P))$$

The difference in proportions is then given by:

$$D = P_B - P_P$$

10. Estimate the standard error of D by:

$$S_D = \sqrt{P_B^2(1-P_B)^2S_B^2 + P_P^2(1-P_P)^2S_P^2 + P_B(1-P_B)P_P(1-P_P)S_R^2 - P_B(1-P_B)P_P(1-P_P)(S_B^2 + S_P^2)}$$

~~The MCMC method for multiple imputation, as previously outlined, Missing data for continuous components of the primary endpoint and binary secondary efficacy endpoints will be imputed using MI as appropriate.~~

~~The above describes the procedure for binary endpoints. For continuous endpoints, the MI procedure will be similar to that described above with the following differences:~~

~~11. The absolute value of the given variable will be imputed. Once imputation has been performed across the 100 iterations specified, any values outside of the range of the given variable will be truncated accordingly.~~

~~12. The change from Baseline values will be computed based on the complete data sets.~~

~~13. The analysis model will be based on ANCOVA (Section 8.3.2 and Section 8.3.3) as opposed to logistic regression.~~

~~For other efficacy variables, MI will be used to account for missing values. The calculation steps impute missing data when possible and where specified. If the imputation model cannot converge, LOCF will be used. The MI procedure will also be similar to that described above will be based on the results provided from the logistic regression model of the multiple imputed datasets. An estimate for the difference in proportions, D, and corresponding standard error, SD, will be computed for each of these datasets. The results from these analyses will be combined into a single estimate of the difference in predicted proportion of response and a 2-sided 95% CI interval using SAS PROC MIANALYZE.~~

~~Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. , for continuous and binary endpoints respectively. However, in the event there are computational challenges with the imputation model (eg, due to a standard deviation of 0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It~~

should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Continuous endpoint

For continuous endpoints (eg, Change from Baseline in DLQI total score at Week 16), the MI method will be applied as follows:

5. The MCMC/monotone regression method described above in Step 1 for binary endpoints will be performed.
6. Based on the multiply imputed data sets obtained for the given variable, the change from Baseline will be derived for each of the 100 complete no inferential statistics will be calculated for the imputed data sets based on the observed Baseline value and the observed/imputed post-Baseline values. Note that if the value itself is being summarized, no additional derivation is needed.
7. If a statistical model is being used for the analysis of the variable, then that will be performed for each imputation in this step. If no statistical model is being used, then simple descriptive statistics will be calculated.
8. For data excluding hs-CRP, the following rules apply. The results of the 100 imputed data sets (based on the statistical model or descriptive statistics) are combined with means and standard errors. Means and standard errors will be calculated using Rubin's rules (via PROC MIANALYZE). **Note that for the calculation of other descriptive statistics such as the median, min, and max, Rubin's rules do not apply. MI estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. Thus, for median, Q1, Q3, minimum and maximum, the following approach will apply:** that will be used when summarizing continuous secondary efficacy variables by subgroup.
 - The data will be summarized by treatment, visit, and imputation, and the summary statistics will be computed.
 - Results will be summarized by taking the mean value of each summary statistic at each visit across all imputations.
 - The number of decimal places will remain the same as the original for display purposes (ie, if the mean was presented to 1 decimal place, then the mean of the means will also be presented to 1 decimal place).

For hs-CRP only, the following rules apply. The hs-CRP data will be presented using the geometric mean, 95% CI for the geometric mean, median, Q1, Q3, minimum and maximum. The change from Baseline will be expressed as the ratio to Baseline in the summaries. The following approach will be applied:

- Following the MI procedure, the ratio to Baseline will be calculated for any of the imputed values
- The natural logarithm of the absolute values and of the ratios to Baseline will be calculated

- The logged values will be summarized (using PROC MEANS) by treatment, visit and imputation
- The datasets will be combined using PROC MIANALYZE in order to get the mean and 95% CI estimates from the absolute values and ratios to Baseline (based on logged data) across imputations
- The estimates of the mean and 95% CI will be back-transformed to obtain the geometric mean and 95% CI on the original scale
- For the median, Q1, Q3, minimum and maximum the procedure outlined above for the other endpoints will be followed

If the imputation model cannot converge, LOCF will be used.

Change #19

Section 4.2.2.2 MI – MCMC/ Referenced-based imputation

The steps of the procedure were updated to:

The steps for the procedure are as follows:

1. For non-monotone (intermittent) missing data, MCMC will be used to impute lesion count data, with Baseline antibiotic use, Hurley Stage at Baseline, and lesion count at Baseline and at each post-Baseline visit (in chronological order) being included in the imputation model. This will be done only once for each participant in order to provide a dataset with monotone missing data.
2. Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcome data at visits $t=1, \dots, T$, where T is Week 16 for HiSCR₅₀.
 - a. *Initialization.* Set $t=1$ (Baseline visit)
 - b. *Iteration.* Set $t=t+1$. Create a data set combining records from bimekizumab- and placebo-treated participants with columns for covariates (Hurley Stage at Baseline and Baseline antibiotic use) and outcomes at visits 1 to t . Outcomes for all bimekizumab-treated participants are set to missing at visit t and set to observed or previously imputed values at visits 1 to $t-1$. Outcomes for placebo-treated participants are set to observed at visit t or observed or previously imputed values at visits 1 to $t-1$. **The outcomes should be sorted in chronological order in the model.**
 - c. *Imputation.* ~~Run MCMC to impute~~ **Impute missing values** for visit t using previous outcomes for visits 1 to $t-1$, Baseline antibiotic use, and Hurley Stage at Baseline. Note that only placebo data will be used to estimate the imputation model since no outcome is available for bimekizumab-treated **participants at visit t . Consequently, the input dataset should include all study participants from placebo but only study participants from the bimekizumab arm that have values at timepoint t missing.**
 - d. Repeat steps 2a-2c, 100 times with different seed values (seeds ranging from 853 to 952) to create 100 imputed complete data sets. **Study participants whose missing values were imputed in the last PROC MI call will be included in the input dataset for the next PROC MI call. Standard rounding rules will be applied to the imputed values. If the MI procedure yields a value outside of**

the pre-defined range for the given variable, the value will be updated to be within the predefined range of values for the endpoint of interest. For example, the imputed value for inflammatory nodules will be updated to 0 in the case of an imputed value less than 0.

- e. *Analysis.* For each completed data set, the outcome (response or change from Baseline) will be calculated using the complete datasets (with no missing data).
3. Each complete imputed data set will then be analyzed based on the statistical model specified in this study (logistic regression). The Week 16 results from logistic regression of each of the 100 imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

Change #20

Section 4.2.2.3 Tipping Point Analysis

The steps for performing the tipping point analysis were updated:

Tipping point analyses will also be implemented as a supportive analysis for the primary efficacy endpoint.

The objective of the tipping point analyses is to **evaluate the sensitivity of results to departures from the missing at random assumption and to identify the point at which departures cause results to "tip" from statistically significant to statistically non-significant.** ~~As such, these identify assumptions about the missing data under which the conclusions from the main analysis change, ie, under which there is no longer evidence of a treatment effect. These tipping point analyses will only be performed if the primary efficacy analysis results in a statistically significant treatment effect ($p < 0.025$). Note that each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo independently for these analyses. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the p-value in this analysis method will be 0.05 instead of 0.025 throughout for that dose.~~

For tipping point analyses, data for participants after As a first step, the intercurrent event date (See Section 3.9) will be changed to missing prior to imputation but will not be changed to non-response after imputation.

The worst-case scenario will be evaluated first. All missing primary endpoint values for (HiSCR₅₀ at Week 16). Specifically, all study participants with a missing HiSCR₅₀ at Week 16 who have been randomized to bimekizumab (where missing values include observations after the intercurrent event date and any other missing values) will be imputed as non-responders, while all missing values for placebo-randomized study participants with a missing HiSCR₅₀ at Week 16 will be imputed as responders. While there is little justification for such an approach, it makes the most putative assumption possible against a bimekizumab treatment effect. After applying this imputation approach, a logistic regression model consistent with the one described for the primary analysis will be applied. If the p-value for the odds ratio of bimekizumab versus placebo remains significant is less than 0.025 for the particular bimekizumab dose regimen, then no further tipping point analyses are needed.

If this analysis based on the worst-case scenario results in a p-value **that is not significant** (eg, greater than 0.025), then additional tipping point analyses will be **performed to identify**. ~~Several assumptions will be made about average outcomes among the point at which results switch or~~ **“tip” from significant to non-significant. Note that subsets of study participants who prematurely discontinued study treatment and hence have a monotone missing data pattern (O’Kelly, 2014). In practice, it implies different delta adjustments will be made to the assumed responses on the monotone missing data in each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo independently for these analyses. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the with various degrees of plausibility with the goal to find for each treatment group the “tipping point” that will significantly reverse the primary result that yielded a p-value in this analysis method will be 0.05 instead of 0.025 throughout for that dose. In the tipping point analysis, a shift parameter or delta adjustment is applied to missing, and subsequently imported primary endpoint values (where missing values include observations after the intercurrent event and any other missing values). These delta adjustments will be done on the lesion count and will be implemented on the primary endpoint as follows:**

- 10. Data after intercurrent event date (See Section 3.9) will be set to missing.**
- 11. The same MCMC method described in Section 4.2.2.1 (Step 1a) will be implemented for non-monotone (intermittent) missing pattern values, using the same imputation model. This will be based on 100 sets of imputations. This will be done only once for each study participant to provide a dataset with monotone missing data.**
- 12. Based on the 100 datasets obtained in Step 2, a monotone regression model will be applied (using the same imputation model as in Step 2) as described in Section 4.2.2.1 (Step 1b).1) while adjusting the imputed values by various delta adjustments. This will be based on 1 imputation.**
- 13. Delta adjustments will be made to imputed lesion count values at Week 16, independently in each treatment group as described below.**
- 14. Delta adjusted imputed values will be truncated so that they are within the range of allowable values for each component.**
- 15. Following the delta adjustments for the individual components lesion counts, of the composite endpoint HiSCR₅₀ will then be derived based on the delta-adjusted multiply imputed data sets obtained for each component endpoint of interest.**

~~Several scenarios will be considered to define these shift parameters. Once defined, the same shift parameter value will be applied on the imputed endpoint value for all visits. Scenario 1 will assume that study participants randomized to bimekizumab and who have missing data have a lower probability of response compared to study participants randomized to placebo with missing data.~~

- ~~— For endpoints for which high scores are associated with a more favorable outcome, it will mean that:~~
 - ~~▪ A negative shift is applied to the imputed value for study participants randomized to bimekizumab to decrease the imputed value.~~

- ~~A positive shift is applied to the imputed value for study participants randomized to placebo to increase the imputed value.~~
- ~~For endpoints for which high scores are associated with a less favorable outcome, it will mean that:~~
 - ~~A positive shift is applied to the imputed value for study participants randomized to bimekizumab to decrease the imputed value.~~
 - ~~A negative shift is applied to the imputed value for study participants randomized to placebo to increase the imputed value.~~

~~For each continuous variable, a set of possible values will be first pre-defined for the shift parameter (example: 0, 1, 2, 3).~~

- 16. Each of the 100 imputed datasets will then be analyzed using a logistic regression model with factors of treatment group, , Baseline Hurley Stage, and Baseline antibiotic use.**
- 17. The results obtained from the 100 logistic regression analyses in Step 7 will be combined for overall inference using Rubin's rules, and the results obtained for each shift parameter will be presented in a single table.**
- 18. Steps 4 to 8 will be repeated so that, at each iteration, missing values are adjusted with a larger delta than at the previous iteration. Depending on the results obtained, shift parameters with more granularity (eg, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9) may also be investigated. The process will go on until the p-value for the odds ratio between bimekizumab and placebo is no longer statistically significant (eg, ≥ 0.025). The odds ratio, 97.5% CI (or 95% depending on the significance level being used for testing), and p-values obtained for each value of delta will be combined in one single table.**

The delta adjustments result in study participants randomized to bimekizumab with missing data having a lower probability of response compared to study participants randomized to placebo with missing data. Since HiSCR₅₀ response is an endpoint for which high lesion counts are associated with a less favorable outcome:

- **A positive adjustment is applied to the imputed value for study participants randomized to bimekizumab in order to increase the imputed value and decrease the likelihood of response.**
- **A negative adjustment is applied to the imputed value for study participants randomized to placebo in order to decrease the imputed value and increase the likelihood of response.**

To start, imputed values within each values within each lesion type, will be adjusted by the same value in each treatment arm. This adjustment will be 5% of the observed range within that lesion type. Depending on the results obtained, this adjustment will be multiplied for step 9 above (2 times, 3 times the initial adjustment) until the p-value is no longer statistically significant. This can be an adjustment of preselected integer values (eg, 1, 2, and 3) or adjustments at intervals equal to a percentage of the allowable range of the component (eg, 5% of range of 10 to give 0.5, 1, 1.5 etc.). Depending on the results obtained, more granular adjustments (eg, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9) may also be implemented to better

~~identify the point at which results "tip". More robust primary analysis results will require larger adjustments to tip the results from significant to insignificant.~~

Additionally, study participants randomized to bimekizumab with an intercurrent event should be set to non-response, after applying the delta adjustment outlined in Step 6 above. This ensures study participants randomized to bimekizumab do not have a higher probability of response in the tipping point analyses compared to the primary analysis (ie, a study participant randomized to bimekizumab who is non-responder in the primary analysis cannot become a responder in the tipping point analyses).

Change #21

New section was added.

Rationale for estimand (Section 4.2.3):

Intercurrent events have been identified within the estimands for this study because their potential to impact efficacy assessments linked with the primary and secondary study objectives. In order to account for the effect of any observed post-randomization intercurrent events on the efficacy analyses, the following estimand strategies will be implemented when evaluating the primary and secondary efficacy endpoints:

- A composite estimand strategy will be used for the primary analysis of the binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, flare),
- A hypothetical estimand will be used for the primary analysis of the continuous secondary endpoints (change from Baseline in DLQI total score and in "worst pain" item for the HSSDD).

Change #22

New section was added.

Composite estimand (Section 4.2.3.1):

A composite estimand strategy as defined in Section 8.2.2 allows incorporation of the two intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) within the definition of the endpoint. These intercurrent events are considered meaningful to the efficacy outcome following receipt of study medication. For example, within the proposed composite estimand framework, a randomized study participant who discontinues from study treatment due to lack of efficacy prior to Week 16 will be considered a treatment failure at Week 16 regardless of the lesion count assessment performed at that visit.

The assumptions and robustness of the primary analysis (modified composite estimand as defined in Section 8.2.2) will be assessed through the sensitivity analyses defined in Section 8.2.3. The impact of intercurrent event handling and data imputation methods on endpoint derivation will also be assessed via the analyses of lesion counts and derived HiSCR variables as specified in Section 8.4.2.1 and Section 8.4.1.1, respectively.

Change #23

New section was added.

Hypothetical estimand (Section 4.2.3.2):

The hypothetical estimand is defined in Section 8.3 and involves a data-driven approach to account for the potential impact of intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) on the analysis of continuous efficacy endpoints. Under this framework, outcomes for study participants without an intercurrent event are analyzed as observed. Conversely, outcomes for study participants with an intercurrent event are imputed via a multiple imputation model, ie any recorded data on or after the intercurrent event will be set to missing and imputed via multiple imputation following the strategy established in Section 4.2.2.1.

Change #24

Section 4.2.4 Dates and times

Partial stop and end date imputation rules were updated:

- Imputation of Partial Start Dates
 - If only the month and year are specified:
 - **If the month and year of first dose of study medication is the same as the month and year of the partial start date, then use the date of first dose of study medication,**
 - **Else, if the month and year of the partial start date are the same as the month and year of a study medication switch date, then use the date of study medication switch,**
 - **Otherwise, use the 1st of the month of the partial start date;**
 - If only the year is specified:
 - **If the year of first dose of study medication is the same as the year of the partial start date, then use the date of first dose of study medication,**
 - **Else, if the year of the partial date is the same as the year of a study medication switch date, then use the date of study medication switch,**
 - **Otherwise, use the 1st of January of the year of the partial start date;**
 - If the start date is completely unknown:
 - **If the stop date is unknown or not prior to the date of first dose of study medication, then use the date of first dose of study medication,**
 - **If the stop date is prior to the date of first dose of study medication, then use the 1st of January of the year of the stop date.**
- Imputation of Partial Stop Dates
 - **If only the month and year are specified, :**
 - **Use the last day of the month of the partial stop date;**
 - **If only the year is specified**
 - **use December 31st of the year of the partial stop date;**

- If the stop date is completely unknown,
 - Do not impute the stop date.

Note that if the stop date or the imputed stop date is prior to the imputed start date, then follow the procedure outlined below:

- If only the year of the start date is specified:
 - If the year of start date is the same as the year of first dose of study medication and the imputed stop date is after the date of first dose of study medication, then set the start date to the date of first dose of study medication,
 - Otherwise, set the 1st January of the year of the start date;
- If only the month and year of start date are specified:
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is on or after the date of first dose of study medication then set the start date to the date of first dose of study medication,
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is before the date of first dose of study medication then set the start date to the 1st of the month of partial start date.

Change #25

Section 4.6 Use of an efficacy subset of participants

The section was updated to:

A sensitivity analysis of the primary endpoint will be performed based on the FAS, the PPS, and the CFS.

Change #26

Section 4.8 Examination of subgroups

This section was updated to:

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, and flare endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period separately. **Additional subgroup analyses will be performed on the change from Baseline in the worst pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.**

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, and flare which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

The following **subgroup variables** will be determined using Baseline data, **except for analgesic use, lesion intervention, and antibody positivity**:

- Age (<40 years, 40 to <65 years, ≥65 years)
- Gender (male, female)

- Disease duration (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for analysis.

- Region (North America [Canada, USA], Western Europe [France, Germany, Ireland, Italy, Spain, United Kingdom], Central/Eastern Europe [Bulgaria, Czech Republic, Hungary, Poland], Asia/Australia [Australia, Israel, Japan])
- Weight (≤100 kg, >100 kg)
- BMI (<25 kg/m², 25 to <30 kg/m², ≥30 kg/m²)
- Race (Black or African American, White, All Other Races [**American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other/Mixed**])
- Systemic antibiotic therapy at randomization (yes, no)
- Prior biologic therapy for any indication (yes, no)
- Prior biologic therapy for HS (yes, no)
- Hurley Stage at Baseline (II or III)
- Analgesic users (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period (Section 6.4.2 specifies how participants are classified as analgesic users)
- Lesion intervention (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period
- Antibody positivity (confirmatory assay: negative or positive)

~~Any analgesic rescue medication taken during the study, lesion intervention (including new post-Baseline antibiotic use or dose adjustments) and antibody positivity are the only subgroups that are not determined by Baseline data. They will be presented in a separate table.~~

Subgroup analyses will also be performed by visit **The following subgroups for analysis on the change from Baseline in the worst pain score as measured by HSSDD and in the DLQI total score through Week 16.** The following subgroups for analysis will be determined based on medication use during the Initial Treatment Period:

- Antihistamines users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users)
- Analgesics users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as analgesic users)
- Systemic antibiotic therapy start/increase after randomization during the Initial Treatment Period (yes, no)

All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

Change #27

Section 5. 1 Study participant disposition

The following sentence was added:

Participants are defined as completing the Initial Treatment Period if they have a Week 16 visit, or if they fail to attend the Week 16 visit but attend at least one visit in the Maintenance Treatment Period.

The following summaries were also added:

To assess participant disposition (entry and periods in the study) during the COVID-19 pandemic, study participant disposition will also be assessed by period of the COVID-19 pandemic (pre – during – post), by comparing the dates of visits (or events) to the dates of the COVID-19 pandemic period. The dates to categorize the periods of the COVID-19 pandemic (pre/during/post) are defined below:

- Pre-COVID-19 pandemic period: Period prior to COVID-19 pandemic start date defined as 11-Mar-2020
- COVID-19 pandemic period: Period from 11-Mar-2020 through the COVID-19 pandemic end date which is currently not defined at the time of approval of the SAP
- Post-COVID-19 pandemic period: Period after the declaration of the end of the pandemic

Change #28

The following new section was added:

Impact of COVID-19 (Section 5.2)

A listing of visits affected by COVID-19 will be presented based on the ES including the visit, date of visit, relationship to COVID-19, impact category and a narrative (short description) of the event. These data will be summarized for non-randomized participants and by treatment group and overall, for enrolled participants.

A summary of study visits by COVID-19 pandemic period (pre – during – post) will be presented for participants enrolled prior to and during the pandemic.

In addition, in order to assess the potential impact of COVID-19 on the collection and reporting of efficacy data, a separate summary on the RS will be presented to display missing data as well as data collected via an alternative modality (e.g.: phone, video call) for efficacy endpoints included in the hierarchy (Section 4.5). For these displays, missing data will be presented only for visits affected by COVID-19, as reported on the dedicated eCRF page. Missing data at other visits and for other reasons will not be included. Note that the remote contingencies for COVID-19 or other exceptional circumstances are not applicable to efficacy assessments and documentation (eg, lesion-based assessments, photography) that require direct face-to-face physician/participant interaction.

Change #29

Section 5.3 Protocol deviations

The following text has been added:

A separate summary of participants with protocol deviations related to COVID-19 will be provided.

A by-participant listing of protocol deviations will be provided. Protocol deviations (eg, missing assessments or visits) related to COVID-19 will be listed separately.

Change #30

Section 6.2 Other Baseline characteristics

The last 5 bullets were updated:

- Duration of disease (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for the summary.

- Baseline antibiotic use (yes, no) (**According to the randomization strata**)
- **Baseline antibiotic use (yes, no) (Derived)**
- Hurley Stage at Baseline (**According to the randomization strata**)
- **Hurley Stage at Baseline (Derived)**

The following text was added or the additional summaries:

In addition, the following Baseline disease characteristics will be summarized by **the derived** Baseline Hurley Stage and by **the derived** Baseline antibiotic use and treatment group for the RS

Change #31

Section 6.4 Prior and concomitant medications

The following sentence was updated:

Prior medications include any medications that started ~~before~~ prior to the start date of study medication. Concomitant medications are **any medication that has a start date on or after the start date of study medication, or any medication that has a start date on or before the last dose of study medication + 28 days (whether placebo or bimekizumab).** ~~medications taken at least 1 day in common with dosing period.~~

The following sentence was added:

Additional summaries for the Initial Treatment Period and Maintenance Treatment Period will be presented for participants taking systemic antibiotic medications that qualify as intercurrent events as described in Section 3.9.

Change #32

Section 8 Efficacy Analyses

This section was updated:

All efficacy analyses of primary, ~~and~~ secondary, **and other** variables will be performed on the RS unless otherwise specified. ~~All efficacy analyses of other efficacy variables will be performed on the RS and MS unless otherwise specified.~~ All efficacy summary tables will be displayed by treatment **sequence** unless otherwise specified. The primary and secondary endpoints, and their components, will also be summarized by **the derived** Hurley Stage at

Baseline (grouping each stage and overall) and treatment **sequence** and by **the derived** Baseline antibiotic use (yes/no and overall) and treatment group.

Change #33

Section 8.2 Primary efficacy endpoint

The following rows in Table 8-2 were added or amended:

Table 8-1: Estimand Details and Attributes for Primary Endpoint

		Estimands for Primary Endpoint			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Primary Objective: To evaluate the efficacy of bimekizumab in study participants with moderate to severe HS					
HiSCR ₅₀	Sensitivity (Section 8.2.3.1)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handles by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.

Table 8-1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.2)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy, whereby all data at and after the intercurrent event will be treated as missing. Composite strategy, as for the primary analysis. Composite strategy, as for the primary analysis.	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – Reference-Based Regression under a missing not at random assumption.
HiSCR ₅₀	Sensitivity (Section 8.2.3.3)	HiSCR ₅₀ response at Week 16	RS	Composite strategy ^a , as for the primary analysis.	A tipping point analysis will be used where various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. The odds ratio versus placebo is based on a logistic regression for each value of delta.

Table 8-1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.8)	HiSCR ₅₀ response at Week 16	CFS	Composite strategy, as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.9)	HiSCR ₅₀ response at Week 16	RS	The same two intercurrent events used for the primary analysis will be used. Any missing data due to COVID-19 will also be considered an intercurrent event. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (nonresponders).	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.10)	HiSCR ₅₀ response at Week 16	RS	Composite strategy, as for the primary analysis.	The odds ratio versus placebo based on a stratified Cochran-Mantel-Haenszel (CMH) test. Missing values not preceded by an intercurrent event will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

CFS=Covid-19 Free Set; CMH=Cochran-Mantel-Haenszel; COVID-19=coronavirus disease 2019; FAS=Full Analysis Set; HiSCR=Hidradenitis Suppurativa Clinical Response; IES=intercurrent event(s) strategy; MCMC=Markov Chain Monte Carlo; MI= multiple imputation; PLS=Population-level summary; Pop=Population; PPS=Per-Protocol Set; RS=Randomized Set

^a The composite estimand strategy will be modified in the tipping point analysis such that participants with intercurrent events will be treated as nonresponders only in the bimekizumab treatment groups.

Change #34

Section 8.2.2 Primary analysis of the primary efficacy endpoint

The following text was updated:

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, and Baseline antibiotic use. The odds ratio versus placebo, p-value (from Wald test), and **97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96 .**

Change #35

New section was added:

Analysis on CFS (Section 8.2.3.8)

The primary efficacy analyses from Section 8.2.2 will be repeated based on the CFS.

Change #36

New section was added:

Analysis including COVID-19 as intercurrent event (Section 8.2.3.9)

An additional sensitivity analysis will include an additional intercurrent event. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and completion of study treatment through Week 16 without receiving systemic antibiotic rescue medication or discontinuing IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for this sensitivity efficacy analysis:

5. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
6. Study participant-level outcome=HiSCR₅₀ at Week 16.
7. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication, discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16, or missing data due to COVID-19. More information is provided in Section 3.9. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, not discontinuing study treatment due to an AE or lack of efficacy through Week 16, and not having missing data due to COVID-19. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation as defined in Section 4.2.1.
8. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

The same imputation techniques and analysis model as in the primary efficacy analyses will then be used.

Change #37

Section 8.3 Secondary efficacy endpoints

The following text has been added:

Sensitivity analyses of the secondary endpoints will be performed on the CFS.

Change #38

Section 8.3.2 Flare by Week 16

The following text was deleted:

Disease flare status will be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in each treatment group. A bar chart of percentage of subjects with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Initial Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Initial Treatment Period will be presented.

Change #39

Section 8.3.3 DLQI Total Score at Week 16

The following paragraph was updated to:

Change from Baseline in DLQI total score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use and Baseline value as a covariate. The least square mean (LSM), standard error (SE), **and 95% CI for the LSM** will be presented by treatment group. For the comparison between placebo and bimekizumab: the difference between the LSM, the associated **97.5% CI** for the contrasts, and the corresponding p-value **will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.**

Change #40

Section 8.3.4 Skin Pain score at Week 16, as assessed by the “worst pain” item in the HSSDD

This section was updated:

The items on the HSSDD assess patients' perception of the core symptoms of HS experienced in the past 24 hours: pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point numeric rating scale (NRS). Two items assess skin pain: worst skin pain and average skin pain.

Weekly averages will be derived for each of the items of the HSSDD **for weeks matching the post-Baseline dosing weeks** up to Week 16. A weekly average is the sum of the scored item over the course of the study week divided by the number of days on which the item was completed. The weekly averages will be relative to the respective visit date except for Baseline, which will be anchored to the first dose of study drug. A weekly average will only be calculated if at least 4 non-missing values (not necessarily consecutive) are available. Otherwise, the HSSDD weekly average for the given question will be set to missing.

Baseline will be computed as the average from the first 7 consecutive day period in which there are at least 4 non-missing entries. That is, first consider the first 7 consecutive days prior to the Baseline visit, but not including the Baseline visit day itself. If there are at least 4 non-missing values (not necessarily consecutive), then the Baseline average will be calculated. If there are less than 4 values, the 7 consecutive day period will move one day earlier. If there are at least 4 non-missing values (not necessarily consecutive) in that period, then the Baseline average will be calculated. This will continue until there are at least 4 non-missing values in a 7 consecutive day period in the 14 days prior to Baseline. If there is no period in which there are at least 4 non-missing entries, then the Baseline value will be set to missing. ~~Baseline will be computed as the average from the 2 weeks prior to Baseline, up to and including the data from the Baseline visit. If less than 7 non-missing values are available for a given question, the Baseline for the given question will be set to missing.~~

Change from Baseline in worst skin pain score is defined as the average Week 16 worst skin pain score minus the Baseline worst skin pain score. Missing data imputation described in Section 4.2.1.2 will be applied to the weekly averages and not to the individual daily PRO data.

Change from Baseline in worst skin pain score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use, analgesic use (Section 6.4.2) and Baseline value as a covariate. A treatment-by-analgesic-use interaction term will also be added to the model and removed if not significant.

The LSM, SE, and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab, the difference between the LSM, the associated 97.5 95% CI for the contrasts, and the corresponding p-value will be presented. **If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose** with a corresponding $Z_{\alpha/2}$ of 1.96.

Change #41

Section 8.4.1.2 Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀

The following text was updated:

Initial Treatment Period

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first HiSCR_{xx} response, Date of Week 16 visit) – Date of **first dose of study medication** Baseline visit + 1, here xx represents 25, 50, 75, 90, 100 respectively. All visits including unscheduled visits are considered.

Participants who discontinue study treatment without achieving a given HiSCR response prior to Week 16 visit will be censored at the date of **last lesion count assessment**. ~~discontinuation.~~

Participants who reach the Week 16 Visit without achieving the given response will be censored at the date of the Week 16 Visit. Participants who experience an intercurrent event **prior to achieving a HiSCR response** will be censored at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

The following text was deleted:

Combined Initial and Maintenance Treatment Period

An additional time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the combined Initial and Maintenance Treatment Period will be calculated as above, where the Week 48 visit is considered instead of Week 16.

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the combined Initial and Maintenance Treatment Period will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group. ~~These summaries will be limited to participants randomized to bimekizumab.~~

Kaplan-Meier plots of time to HiSCR responses will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.

Maintenance Treatment Period

~~For participants randomized to placebo, an additional time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the Maintenance Treatment Period will be calculated as:~~

~~Min (Date of first HiSCR_{xx} response, Date of Week 48 visit) — Date of Week 16 visit + 1, here xx represents 25, 50, 75, 90, 100 respectively. All visits including unscheduled visits are considered.~~

~~Participants who discontinue study treatment without achieving a given HiSCR response prior to Week 48 visit will be censored at the date of discontinuation. Participants who reach the Week 48 Visit without achieving the given response will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event will be censored at the date of the intercurrent event. Participants will be censored at Week 16 if there is no Post Week 16 lesion count assessment.~~

~~Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period will each be estimated and presented using the Kaplan-Meier product limit method for the placebo/bimekizumab 320mg Q2W treatment group.~~

~~The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.~~

Change #42

New section was added:

HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ among Week 16 Responders (Section 8.4.1.4)

See Section 8.4.1.1 for the derivation of HiSCR response.

Summaries of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at each visit from Week 16 through Week 48 will be summarized based on a subset of participants in the MS who achieve response at Week 16. The summaries will be as follows:

- HiSCR₅₀ responder rate based on participants who achieved HiSCR₅₀ response at Week 16
- HiSCR₇₅ responder rate based on participants who achieved HiSCR₇₅ response at Week 16
- HiSCR₉₀ responder rate based on participants who achieved HiSCR₉₀ response at Week 16
- HiSCR₁₀₀ responder rate based on participants who achieved HiSCR₁₀₀ response at Week 16

Line plots of the above HiSCR responder rate categories over time (from Week 16 to Week 48), by treatment group, will be produced.

Change #43

Section 8.4.1.5 Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders

The following text has been updated:

See Section 8.4.1.1 for the derivation of HiSCR response.

Time to loss of response will be based on the MS and include only participants who had the corresponding HiSCR response at Week 16 (considering intercurrent event handling from the composite estimand described in Section 8.2.2).

Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) is defined as: Date of loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ - Date of Week 16 treatment **administration** + 1.

Time to loss of response will be estimated and presented using the Kaplan-Meier product-limit method for each treatment. Participants who experience an intercurrent event **prior to loss of response** will be considered as having lost response on the date of intercurrent event.

Participants who reach the Week 48 Visit without loss of response will be censored at the date of the Week 48 Visit. Participants who discontinue treatment or study, for reasons other than those already defined for an intercurrent event, and who have not yet displayed loss of response by the time of withdrawal, will be censored at the date of **the last lesion count assessment withdrawal**.

~~The summary for each HiSCR response will include only participants who had the corresponding HiSCR response at Week 16.~~

Change #44

Section 8.4.1.6 Partial Response

The following paragraph was updated:

The number and percentage of participants who are partial responders at Week 16 and become HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders in the Maintenance Treatment Period will be summarized by treatment group and visit. These analyses will be based on the subset of participants in the **MS RS** that are partial responders but not HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders, respectively, at Week 16. **These summaries will be based on observed case data and will not consider the occurrence of intercurrent events.**

Change #45

New section was added:

Flare relative to Baseline (Section 8.4.3)

See Section 8.3.2 for the derivation of flare.

Disease flare status will be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in each treatment group. This summary will also include the number of participants with any flare in the Initial Period, Maintenance Period, and the combined Initial and Maintenance Period. A bar chart of percentage of participants with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Initial Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Initial Treatment Period will be presented.

Change #46

The following section was deleted:

Flare by Week 48 (Section 8.4.3)

See Section 8.2.1 for the derivation of AN count.

Disease flare by Week 48 is defined when at least a 25% increase in AN count with an absolute increase of ≥ 2 AN relative to Week 16 is observed by Week 48. A participant's disease flare status (yes/no) will be determined at each visit in the Maintenance Treatment Period using these criteria and will be listed with the other lesion count assessment data in the data listings.

The number of participants who experience at least 1 disease flare by Week 48 will be summarized by treatment group.

Disease flare status during the Maintenance Period will also be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in the Maintenance Treatment Period in each treatment group. A bar chart of percentage of subjects with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Maintenance Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Maintenance Treatment Period will be presented.

Change #47

Section 8.4.4 Time to flare by Week 16

This section was updated as follows:

See Section 8.3.2 for the derivation of flare ~~by Week~~.

Time to flare (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first flare, Date of Week 16 visit) – Date of **first dose of study medication**
~~Baseline visit~~ + 1. All visits in the Initial Treatment Period including unscheduled visits are considered.

Participants who discontinue study treatment without experiencing a flare prior to Week 16 Visit will be censored at the date of **last lesion count assessment**. ~~discontinuation~~. Participants who reach the Week 16 Visit without experiencing a flare will be censored at the date of the ~~the~~ Week 16 Visit. Participants who experience an intercurrent event **prior to experiencing a flare** will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.

The median time to **flare response**, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

Change #48

Section 8.4.5 Time to flare by Week 48

This section was updated as follows:

See Section 8.3.2 for the derivation of flare **relative to Baseline** by Week 48.

Maintenance Treatment Period

~~Time to flare (in days) during the Maintenance Treatment Period will each be calculated as:~~

~~Min (Date of first flare, Date of Week 48 visit) – Date of Week 16 visit + 1. All visits in the Maintenance Treatment Period including unscheduled visits are considered.~~

~~Participants who discontinue study treatment without experiencing a flare prior to Week 48 visit will be censored at the date of discontinuation. Participants who reach the Week 48 Visit without experiencing a flare will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Week 16 if there is no Baseline lesion count assessment or no Post Week 16 lesion count assessment.~~

~~Time to flare will each be estimated and presented using the Kaplan-Meier product limit method for each treatment group.~~

~~Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.~~

~~The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.~~

Combined Initial and Maintenance Treatment Period

Time to flare (in days) during the combined Initial and Maintenance Treatment Period will be calculated as:

Min (Date of first flare, Date of Week 48 visit) – Date of **first dose of study medication** ~~Baseline visit~~ + 1. All visits in the up to Week 48 including unscheduled visits are considered.

Flare will be defined relative to the Baseline visit. Participants who discontinue study treatment without experiencing a flare prior to Week 48 visit will be censored at the date of **last lesion count assessment**. ~~discontinuation~~. Participants who reach the Week 48 Visit without experiencing a flare will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event **prior to experiencing a flare** will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group. ~~This summary will be limited to participants randomized to bimekizumab.~~

~~Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.~~

The median time to **flare response**, including the 2-sided 95% confidence interval, will be calculated for each treatment.

Change #49

Section 8.4.6 International Hidradenitis Suppurativa Severity Score System (IHS4)

The following sentence was added:

The observed IHS4 score, change and percentage change from Baseline will be summarized by treatment group and visit. **Missing IHS4 scores will be imputed using the multiple imputation procedure specified in Section 4.2.2.1, where IHS4 scores will be derived based on the imputed lesion counts.**

Change #50

Section 8.4.10 Time to initiation of systemic rescue therapy in the Initial Treatment Period

The following text was updated:

See Section 3.9 for the definition of a systemic antibiotic rescue therapy.

Time to initiation of systemic rescue therapy (in days) during the Initial Treatment Period will be calculated as:

Min (Date of initiation of rescue therapy, Date of change in the dose/type of current antibiotic, Date of Week 16 visit) – Date of **first dose of study medication** ~~Baseline visit~~ + 1.

Participants who discontinue ~~the study treatment~~ without initiating systemic rescue therapy prior to Week 16 visit will be censored at the date of discontinuation. Participants who reach the Week 16 Visit without initiating systemic rescue therapy will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no ~~Baseline lesion count assessment or no Post-Baseline visit lesion count assessment~~.

Change #51

Section 8.4.11 Time to an intercurrent event in the Initial Treatment Period

The following text was updated:

See Section 3.9 for the definition of an intercurrent event.

Time to an intercurrent event (in days) during the Initial Treatment Period will be calculated as:

Min (Date of **intercurrent event** initiation of rescue therapy, Date of change in the dose/type of current antibiotic, Date of withdrawal due to AE or lack of efficacy, Date of Week 16 visit) – Date of **first dose of study medication** Baseline visit + 1.

Participants who discontinue the study treatment without experiencing an intercurrent event prior to Week 16 visit will be censored at the date of discontinuation. **That includes participants who discontinue from the study for reasons other than Adverse Event and Lack of Efficacy.** Participants who reach the Week 16 Visit without experiencing an intercurrent event will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline visit lesion count assessment.

Change #51

Section 8.4.12 Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)

The following sentence was updated:

Change from Baseline in Worst Pain score and Worst Itch score at Week 16 will additionally be summarized **by visit and** by analgesic and antihistamine use status (Section 6.4.2), respectively.

Change #53

Section 8.4.13 Hidradenitis Suppurativa Symptom Questionnaire (HSSQ)

The following text was updated:

HSSQ response for pain **item score** is defined as at least a 30% reduction and at least a 1-unit reduction from Baseline in HSSQ among study participants with a score of ≥ 3 at Baseline.

The number and percentage of responders for pain **item score** will be summarized by treatment group and visit **based on the MS.**

The number and percentage of participants who were responders at any timepoint in the Maintenance Treatment Period will be summarized by treatment group for the skin pain score **based on the MS.**

Change from Baseline in pain score and itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

Change #54

Section 8.4.15 Hidradenitis Suppurativa Quality of Life (HiSQOL)

The following text was updated:

Summary statistics of the actual values and change from Baseline values will be used to summarize HiSQOL **domain and total scores** for each visit by treatment group. The table will

display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3. **The imputed HiSQOL total score will be derived based on the imputed subscales.**

The number and percentage of participants that complete the HiSQOL will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit.

A by-participant listing of the HiSQOL questionnaire, HiSQOL responses, **domain and total scores** and change from Baseline will be provided.

Change #55

Section 9.1 Pharmacokinetics

The following sentence was updated:

~~All However, all~~ PK concentrations **collected will be listed irrespective of the dosing or sampling occurring out of window.**

Change #56

Section 8.2.3.1 Derivation of palmoplantar IGA response

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-tiered assay approach: screening assay, confirmatory assay, and titration assay. Samples confirmed as positive within the confirmatory assay will be further evaluated **for the presence of in a neutralizing anti-assay to evaluate the potential of the ADAb to neutralize the activity of bimekizumab antibodies specific to IL-17AA, IL-17FF (IL-17A or IL-17F, or both) in vitro.** Samples will be taken at Baseline, then at study Weeks 4, 8, 12, 16, 20, 24, 36 and 48, and at PEOT and SFU timepoints.

~~The screening~~ Screening, confirmatory, and titer cut **point will be used to determine points of the status of anti-bimekizumab antibodies in respective assays will be determined by the test sample as Positive Screen (PS) or Negative Screen (NS). For bioanalytical laboratory based either on commercially available drug naïve samples presenting anti-bimekizumab antibody levels that are PS, a further confirmatory assay will be performed, and the result of which will be reported as either Positive Immunodepletion (PI) or Negative Immunodepletion (NI).**

ADAb status for each sample will be derived as follows:

- **Sample values that are either NS, or PS and NI and where the bimekizumab concentration is less than the validated ADAb assay drug tolerance limit will be defined as anti-bimekizumab antibody negative.**
- **Sample values that are either NS, or PS and NI and where the bimekizumab concentration exceeds the validated ADAb assay drug tolerance limit will be defined as inconclusive.**

- **Sample values that are PS and PI will be defined as ADA b positive (regardless of availability of a titer value)**
- **Missing or non-evaluable samples will be defined as missing**

Positive immunodepletion samples will be titrated, and the ADA b titer (reciprocal dilution factor including minimum required dilution) will be reported. Subsequently, PI samples will also be subject to a neutralizing assay to evaluate the potential of ADA b to neutralize the target binding of bimekizumab (IL-17AA or IL-17FFIL17F or both) in vitro.

The following definitions will be applied regarding ADA b status of each test samples:

- ~~An ADA b status will be confirmed as positive for any sample with an ADA b level that is positive screen and positive immunodepletion.~~
- ~~An ADA b status of negative will be concluded for any sample with an ADA b level that is either negative screen or (positive screen and negative immunodepletion).~~

~~If the titer for an ADA b level that is positive screen and positive immunodepletion is missing, then a conservative approach will be used and ADA b status will be consider as positive. No imputation rules apply for the missing titer. If the ADA b level is positive screen but no confirmatory result could be determined, then a conservative approach will be used and ADA b status will be consider as positive.~~

~~Anomalous values will be not included in summaries/analysis and will be reviewed and flagged by the Clinical Pharmacologist.~~

There are different levels of classification for ADA b status, the definitions are as follows:

~~For each participant an overall ADA b status will be derived:~~

- ~~Overall Positive is defined as having at least one value that is confirmed positive during the treatment period.~~
- ~~Overall Negative is defined as having no values that are confirmed positive at any time in the treatment period.~~

~~The treatment period does not include Baseline/pre-treatment samples or SFU.~~

~~Furthermore, the following subcategories for each subject will be derived:~~

- **Pre ADA b negative – treatment-emergent ADA b negative (Category 1): Includes study participants who are anti-bimekizumab antibody negative at Baseline and anti-bimekizumab antibody negative at all sampling points during the period of interest (one post-Baseline missing/inconclusive sample is allowed for subjects with pre- anti-bimekizumab antibody negative sample). This group also includes study participants who have a missing or inconclusive sample (either missing or inconclusive or insufficient volume) at Baseline (ie, pre-treatment) with all post-Baseline samples as ADA b negative.**
- **Pre ADA b negative – treatment-emergent ADA b positive (Category 2): Includes study participants who are ADA b negative at Baseline and ADA b positive at any sampling points post-Baseline during the period of interest. This group also includes study**

participants who have a missing sample (either missing or insufficient volume) at Baseline (ie, pre-treatment) with 1 or more post-Baseline samples as ADAb positive.

- **Pre ADAb positive – treatment-emergent reduced ADAb (Category 3):** Includes study participants who are ADAb positive at Baseline, and ADAb negative at all sampling points post-Baseline during the period of interest.
- **Pre ADAb positive – treatment-emergent unaffected ADAb positive (Category 4):** Includes study participants who are ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with titer values of the same magnitude as Baseline (ie, less than a predefined fold difference from the Baseline titer).
- **For this analysis, this is set at an increase of less than the validated Minimum Significant Ratio (MSR) of 2.07-fold from Baseline.**
- **Pre ADAb positive – treatment-emergent ADAb boosted positive (Category 5):** Includes study participants who ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with increased titer values compared to Baseline (equal to or greater than a predefined fold difference increase from Baseline titer which will be defined within the validation of the assay).
 - **For this analysis, this is set at an increase equal to or greater than the validated MSR of 2.07-fold from Baseline.**
 - **Note:** for any study participant who is ADAb positive at Baseline and ADAb positive at a post-Baseline time point during the period of interest, but for whom titers are not available to determine treatment unaffected or treatment boosted status, the study participant will be considered as treatment boosted, assuming no other samples are available.
- **ADAb Inconclusive (Category 6):** Includes study participants who have an ADAb positive Baseline (pre-treatment) sample and some post-Baseline samples during the period of interest are missing or inconclusive, while other post-Baseline samples are ADAb negative.
- **Total treatment-emergent ADAb positivity (Category 7 [Categories 2 and 5 combined]):** Includes study participants who are pre ADAb negative – treatment-emergent ADAb positive (Category 2) and pre ADAb positive – treatment boosted ADAb positive (Category 5).
- **Total prevalence of pre- ADAb positivity (Category 8 [Categories 3, 4, 5 and 6 combined]):** Study participants that are tested ADAb positive at Baseline.
- **Missing:** Includes study participants who are ADAb negative, missing, or inconclusive at Baseline with some post-Baseline samples as missing or inconclusive, and other samples as ADAb negative.

Derivation for above classification will be different for the interim analysis and the final analysis. SFU data will be considered only for the final data analysis. That is, each instance of “excluding SFU” in the categories above, should be changed to “including SFU.”

In the case that a sample is collected 1 or more days following the scheduled visit date in which the drug was administered, the **ADAb anti-bimekizumab antibody** results for that sample will be associated with the scheduled visit and summarized accordingly. Such samples will also be considered when **ADAb anti-bimekizumab antibody** results are summarized over a given study period.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by participants. All analyses will be run on the **AMS Active Medication Set**, unless specified otherwise.

- Summary of **ADAb anti-bimekizumab antibody** status overall and by each visit separated by treatment group
- Summary of the time-point of the first occurrence of **ADAb anti-bimekizumab antibody** positivity during the treatment period by treatment group. A plot of the titer by time to first **ADAb anti-bimekizumab antibody** positivity will be prepared.
- All individual participant-level **ADAb anti-bimekizumab antibody** results will be listed.
- The number and percentage of participants in each of the **8 ADAb anti-bimekizumab antibody** categories during the treatment period by treatment group, with an additional category combining participants in categories 2 and 5, summarized as total treatment emergent. In addition, the count and percentage of participants who are pre anti-bimekizumab positive will be calculated (this is the sum of categories 3, 4, and 5).
- The prevalence of immunogenicity, separated by treatment group, and defined subcategory, will be reported by visit, defined as (cumulative) proportion of participants having confirmed positive **ADAb anti-bimekizumab antibody** samples at any visit up to and including that visit. Missing samples will not be included in the denominator.
- The time to achieving treatment-emergent **ADAb anti-bimekizumab antibody** positivity, separated by treatment group and **defined subcategory** sub-categories 2 & 5 above, will be analyzed based on Kaplan-Meier methods. Participants will be considered to have an event at the time point at which treatment emergent **ADAb anti-bimekizumab antibody** positive is first achieved (taking the MSR into consideration for sub-category 5). Participants classified as treatment-emergent **ADAb anti-bimekizumab antibody** negative will be censored at the time of the last available **ADAb anti-bimekizumab antibody** result.
- A summary of HiSCR₅₀ responders at Week 16, separated by treatment group, as a function of ADAb titer will be presented graphically. This will be repeated for HiSCR₇₅ responders.
- Individual plots of plasma bimekizumab concentrations/ **ADAb anti-bimekizumab antibody** titer both plotted on the Y-axes by visit (x-axis) for the full treatment period (excluding SFU for interim analyses and including SFU for final analyses) will be presented for participants with and without HiSCR₅₀ response at Week 16.
- Spaghetti plots of ADAb titer (y-axis) by visit (x-axis), separated by treatment group for all **ADAb anti-bimekizumab antibody** positive participants, including Baseline positive participants.

- Box plots of ADAb titer (logscale) by time to first ADAb positivity by treatment group.

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, 2 categories will be used:

- **ADAb anti-bimekizumab antibody positive** – This is defined as participants who have **ADAb anti-bimekizumab antibody** levels above the specified cut point on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- **ADAb anti-bimekizumab antibody negative** – Participants who are not defined as anti-bimekizumab positive (as described above) will be defined as **ADAb anti-bimekizumab antibody negative**.

The groups for defining **ADAb anti-bimekizumab antibody** status for safety subgroup analyses are as follows:

- AEs starting before first **ADAb anti-bimekizumab antibody** positive result
- AEs starting on or after first **ADAb anti-bimekizumab antibody** positive result
- AEs for participants who were always **ADAb anti-bimekizumab antibody** negative

Change #57

Section 10.1.1 Exposure during the Initial Treatment Period

This section was split into 2 subsection 10.1.1.1 and 10.1.1.2 for exposure duration (days) and time at risk (days), respectively.

Change #58

Section 10.1.1.1 Study medication duration (days)

The following text was updated:

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Period + 14 days.

Note: The use of 14 days assumes a Q2W dosing interval (bimekizumab 320mg Q2W and placebo). For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Initial Treatment Period – Date of first dose in the Initial Period + 28 days).

Note: If date of last dose in the Initial Treatment Period + 14 days (or ~~date of last bimekizumab dose in the Initial Treatment Period~~ + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Treatment Period + 1.

Change #59

Section 10.1.1.2 Time at risk (days)

The section was updated:

Definitions for time at risk (days) are provided as follows:

~~For participants who permanently discontinue study treatment:~~

- ~~• Date of last dose – date of first dose + 14 days~~

~~The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Initial Treatment Period – date of first dose in the Initial Period + 28 days).~~

~~Note: If date of last dose + 14 days (or date of last dose of bimekizumab + 28 days for Q4W dosing) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:~~

~~— Final visit date (including PEOT, but not including SFU) – date of first dose + 1.~~

Time at risk (days)

- For participants who complete the Week 16 visit and continue to the Maintenance Treatment Period:
 - Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Period + 1.
- For participants who discontinue on or prior to the final visit of the Initial Period, use the minimum of the following:
 - **Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Treatment Period + 141**
 - The total number of days in the Initial Treatment Period (112 days). For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE, but will be excluded from the output based on the Initial Treatment Period. However, these AEs will be included in the AE summaries for Maintenance Treatment Period.
 - Date of last clinical contact – Date of first dose in the Initial Treatment Period + 1.
- For participants who die prior to the final visit of the Initial Treatment Period: Date of death – date of first dose in the Initial Period + 1.

Change #60

Section 10.1.2 Exposure during the Maintenance Treatment Period

This section was split into 2 subsection 10.1.2.1 and 10.1.2.2 for exposure duration (days) and time at risk (days), respectively.

Change #61

Section 10.1.2.1 Study medication duration (days)

The section was updated:

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 14 days.

The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W in the Maintenance Treatment Period, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 28 days).

Note: If date of last dose in the Maintenance Treatment Period + 14 days (or ~~date of last bimekizumab dose in the Maintenance Treatment Period~~ + 28 days in the case of Q4W dosing) extends to a date beyond the final visit date of the Maintenance Treatment Period (not including SFU), then this calculation reverts to:

- Final visit date of the Maintenance Treatment Period (not including SFU) – date of first dose in the Maintenance Treatment Period + 1.
- ~~Note:~~ For participants who die during the Maintenance Treatment Period, then this calculation reverts to:
 - Date of death – Date of first dose in the Maintenance Treatment Period + 1.

~~For participants who permanently discontinue study treatment:~~

- ~~Date of last dose – date of first dose + 14 days~~

~~The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W in the Maintenance Treatment Period, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 28 days).~~

~~Note: If date of last dose + 14 days (or date of last dose of bimekizumab + 28 days for Q4W dosing) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:~~

- ~~– Final visit date (including PEOT, but not including SFU) – date of first dose + 1.~~

Change #62

Section 10.1.2.2 Time at risk (days)

The text was updated:

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study): ~~Final visit date of the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 1.~~
 - **Date of last visit of the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + 1.**
- For participants who die prior to the final visit of the Maintenance Treatment Period: ~~Date of death – date of first dose in the Maintenance Period + 1.~~
 - **Date of death – Date of first dose in the Maintenance Period + 1.**

- For all other participants, use the minimum of the following:
 - Date of last dose in the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + ~~141~~40 days.

Change #63

Section 10.1.3 Exposure during the Initial and Maintenance Treatment Period

This section was split into 2 subsection 10.1.3.1 and 10.1.3.2 for exposure duration (days) and time at risk (days), respectively.

Change #64

Section 10.1.3.1 Study medication duration (days)

The section was updated:

Definitions for study medication duration (days) are provided as follows:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Period.

Note: The algorithms for calculating these durations are specified in Section 10.1.1.1 and Section 10.1.2.1.

Note: If date of last dose in the Initial Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Periods - 1.

~~For participants who do not switch study treatments:~~

- ~~• Date of last dose – Date of first dose + 14 days.~~

~~The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose – date of first dose + 28 days).~~

~~Note: If date of last dose + 14 days (or date of last bimekizumab dose in the Maintenance Treatment Period + 28 days in the case of Q4W dosing) extends to a date beyond the final visit date (including PEOT, not including SFU), then this calculation reverts to:~~

- ~~– Final visit date (including PEOT, not including SFU) – Date of first dose + 1.~~

- ~~• For participants who die, if date of last dose + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of death, then this calculation reverts to:~~

- ~~– Date of death – Date of first dose + 1.~~

~~For participants who switch study treatments (between Initial and Maintenance Treatment Periods):~~

- ~~• Initial Treatment Period (attributed to initially randomized treatment):~~

- ~~– Date of last dose in the Initial Period – Date of first dose in the Initial Period + 14 days.~~

~~Note: Participants who switch study treatments are on a Q2W dosing schedule for the Initial Treatment Period.~~

~~Note: If date of last dose in the Initial Treatment Period + 14 days extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:~~

~~— Date of first dose in the Maintenance Treatment Period — Date of first dose in the Initial Period + 1.~~

- ~~• Maintenance Treatment Period (attributed to the treatment initiated in the Maintenance Treatment Period):~~

~~— Use the study medication duration algorithm specified for the Maintenance Treatment Period in Section 10.1.2.1.~~

Change #65

Section 10.1.3.2 Time at risk (days)

This section was updated:

~~For participants who do not switch study treatments:~~ Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study):
 - Final visit date – Date of first dose + 1.
- For participants who die prior to the final visit:
 - Date of death – Date of first dose in the + 1.
- For all other participants, use the minimum of the following:
 - Date of last dose – Date of first dose + 141 days.
 - Date of last clinical contact – Date of first dose + 1.

Note: This group could include participants who discontinue early, participants who complete the Maintenance Treatment Period as scheduled but choose not to continue into an extension study, or participants who are ongoing in the SFU period at the time of the data snapshot (in the case of the interim analysis).

~~For participants who switch study treatments (between Initial and Maintenance Treatment Periods):~~

- ~~• Initial Treatment Period (attributed to initially randomized treatment):~~

~~— Date of first dose in the Maintenance Treatment Period — Date of first dose in the Initial Period + 1.~~

~~(Note: This assumes that anyone in this category has completed the Initial Treatment Period and doses [with a new study treatment] in the Maintenance Treatment Period.)~~

- ~~• Maintenance Treatment Period (attributed to the treatment initiated in the Maintenance Treatment Period):~~

~~Use the time at risk algorithm specified for the Maintenance Treatment Period in Section 10.1.2.2.~~

Change #66

Section 10.2.1 Data considerations

The following sentence was added:

If the intensity of an AE is unknown, it is considered as severe. If the relationship to study drug is missing, it is considered as related. **Note that if the seriousness of an adverse event is unknown, every attempt should be made to resolve this prior to a snapshot for an interim analysis or database lock; in the exceptional case that the seriousness of an adverse event is still missing then no imputation should be applied for this characteristic.**

Change #67

Section 10.2.2 AE summaries

The following text was deleted:

The following summaries will be provided by treatment group for the Initial Treatment Period, ~~Maintenance Treatment Period~~, and the Initial and Maintenance Treatment Period combined based on the SS, ~~MS~~, and AMS respectively.

The following summaries were added:

- Incidence of TEAEs – Suspected and Confirmed COVID-19 cases by SOC, HLT and PT

Suspected and confirmed COVID-19 cases will be identified with the preferred terms “Corona virus infection” or “Corona virus test positive”.

The following subset of tables will also be presented for the Maintenance Treatment Period using the MS:

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT

Change #68

Section 10.2.3.1 Infections (serious, opportunistic, fungal and TB)

The following text was updated:

- **Incidence of Fungal Infection TEAEs per 100 subject years by SOC, HLT and PT**

Fungal infections will be summarized in a stand-alone table. The table will include all TEAEs (serious and non-serious) that code into the **High Level Group Term (HLGT)** ~~HLT~~ “Fungal infectious disorders”

- **Incidence of Opportunistic Infection TEAEs per 100 subject years by SOC, HLT and PT**

Opportunistic infections (including tuberculosis) will be summarized in a stand-alone table. The table will include all TEAEs identified using UCB-defined search criteria. (refer to Excel spreadsheet on “OI – MedDRA v19.0.xlsx” in “Bimekizumab Safety Topics of Interest.docx”).

Change #69

Section 10.2.3.3 Major adverse cardiac event

The following sentence was added:

A separate table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type (24 total), the individual PTs that fall within each event type will be summarized. **The other 10 MACE events not listed in the table are described in the adjudication committee charter.**

Change #70

Section 10.2.3.5 Suicidal ideation and behaviors

The following paragraph was added:

A separate table will present the adjudicated neuropsychiatric events by type. For each neuropsychiatric event type (6 total), the individual PTs which fall within each event type will be summarized. It will include events adjudicated as SIB and events adjudicated as non-suicidal. Note that the event type Suicidal ideation can be classified as either SIB or non-suicidal.

Change #71

Section 10.2.3.6 Inflammatory bowel disease

This section was updated:

An external inflammatory bowel disease (IBD) adjudication committee will evaluate potential IBD events and will classify each one as follows:

- Event Type Code 1: Possible IBD – Crohn’s Disease
- Event Type Code 2: Probable IBD – Crohn’s Disease
- Event Type Code 3: Definite IBD – Crohn’s Disease
- Event Type Code 4: Possible IBD – Ulcerative Colitis
- Event Type Code 5: Probable IBD – Ulcerative Colitis
- Event Type Code 6: Definite IBD – Ulcerative Colitis
- Event Type Code 7: Possible IBD – Unclassified
- Event Type Code 8: Probable IBD – Unclassified
- Event Type Code 9: Definite IBD – Unclassified
- Event Type Code 10: Symptoms not consistent with IBD
- **Event Type Code 11: Possible Inflammatory Bowel Disease – Microscopic Colitis**
- **Event Type Code 12: Probable Inflammatory Bowel Disease – Microscopic Colitis**

- **Event Type Code 13: Definite Inflammatory Bowel Disease – Microscopic Colitis**
- **Event Type Code 14: Possible Inflammatory Bowel Disease – no further differentiation possible**
- **Event Type Code 15: Probable Inflammatory Bowel Disease – no further differentiation possible**
- **Event Type Code 16: Definite Inflammatory Bowel Disease – no further differentiation possible**
- **Event Type Code 99: Not enough information to adjudicate**

A table for adjudicated ~~definite~~ IBD events (event type codes **1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15** and **169**) as determined by the adjudication committee will be produced. **It will summarize events determined by the adjudication committee as definite IBD (event type codes 3, 6, 9, 13, and 16), probable IBD (event type codes 2, 5, 8, 12, and 15) and possible IBD (event type codes 1, 4, 7, 11, and 14). Definite and probable IBD will also be aggregated and summarized.** This table will be ~~produced overall~~, as well as stratified by subjects with or without a previous medical history of IBD. Previous medical history of IBD will be determined using the information recorded on the **History of IBD** Extra-Articular Assessment at Screening CRF page ("Does subject have a history of IBD?").

A ~~separate table will present the for~~ adjudicated gastrointestinal ~~probable IBD events by type.~~ **For each gastrointestinal event type (17 total), the individual PTs which fall within each event type will be summarized. It will include events codes 2, 5, and 8) as determined by the adjudication committee as definite IBD probable IBD and will be produced.** This table will be ~~produced overall~~, as well as stratified by participants with or without a previous medical history of IBD.

A table for adjudicated ~~possible IBD. It~~ events (event type codes ~~1, 4, and 7~~) as determined by the adjudication committee will be produced. This table **will also include events determined as Symptoms not consistent with** be produced overall, as well as stratified by participants with or without a previous medical history of **IBD (event type code 10) and Not enough information to adjudicated (event type code 99).**

A listing of all events identified for potential review by the IBD adjudication committee will be produced. This listing will indicate whether each event was escalated to the committee for formal review/adjudication.

A ~~separate table and listing~~ will present the adjudicated IBD events by type. For each IBD event type (event type codes 1 through ~~1640~~ and 99; ~~1744~~-total), the individual PTs which fall within each event type will be **listedsummarized.**

A third listing will present the individual diagnostic criteria met for each adjudicated IBD event.

Change #72

Section 10.2.3.7 Hypersensitivity (including anaphylaxis)

The following text was updated:

A separate table will be prepared to summarize hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. **In addition, a separate table will be prepared to summarize serious hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All serious TEAEs which code to a PT included in the Scope=Narrow search will be included in this table.** An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

Furthermore, **a separate table will be prepared to summarize injection site reactions, identified using the HLTs: “Administration site reactions NEC” and “Injection site reactions”.**

Change #73

Section 10.3 Clinical laboratory evaluations

The following text was added:

For tables where data are summarized by visit, unscheduled and repeat visits will not be summarized, but these data will be included in listings. For tables where multiple measurements over a period of time are considered (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (**values observed more than 140 days after the last administration of study medication are not considered**). All summaries will be presented in SI units and will be based on observed case values.

CTCAE grading was updated:

Markedly abnormal values are defined as those with a severity of Grade 3 and above based on the CTCAE criteria **Version 4.03**, (U.S. Department of Health and Human Services 2017).

And Table 10-2 was updated:

Table 10–2: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Creatinine ¹ Creatinine	mg/dL	> 3.0 x Baseline or >3.0 x ULN	umol/mole μmol/L	> 3.0 x Baseline or >3.0 x ULN	AH
Glucose	mg/dL	<40 >250	mmol/L	<1.7	AL
				>13.9	AH
Calcium	mg/dL	>12.5 <7.0	mmol/L	>3.1	AH
				<1.75	AL
Magnesium	mg/dL	>3.0 <0.9	mmol/L	>1.23	AH
				<0.4	AL
Potassium	mmol/L	>6.0 <3.0	mmol/L	>6.0	AH
				<3.0	AL

Table 10–2: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Sodium	mmol/L	>155 <130	mmol/L	>155 <130	AH AL
Cholesterol	mg/dL	>400	mmol/L	>10.34	AH

1 The markedly abnormal definitions for creatinine are based on the logical or, if either criterion is met the creatinine value will be designated as abnormal high.

And the following text was added:

- **Total Bilirubin:** >1.5xULN, >2xULN
- **ALP:** >1.5xULN

For any participant with at least one markedly abnormal LFT (AST >3xULN, ALT >3xULN, bilirubin >3xULN, or ALP >1.5xULN) the New Ratio (nR) will be calculated as the ratio of either ALT or AST (whichever is higher) to ALP, all expressed as multiples of their ULN as follows:

- $nR = [\text{maximum}(\text{AST}/\text{ULN or ALT}/\text{ULN})]/(\text{ALP}/\text{ULN})$

Any pDILI will be summarized (all criteria must be met at the same assessment):

- (AST or ALT > 3xULN) and Total Bilirubin > 1.5xULN
- (AST or ALT > 3xULN) and Total Bilirubin > 2xULN

In addition, a table will be produced to summarize potential Hy's Law cases. The following definition will be used in that table:

- $[\text{AST} \geq 3\text{xULN or ALT} \geq 3\text{xULN}]$ and Total Bilirubin $\geq 2\text{xULN}$ in the absence of ALP $\geq 2\text{xULN}$

In order to meet the above **potential Hy's Law** criteria, a participant must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation, if applicable) at the same assessment. For example, a participant who experiences a ≥ 2 x ULN elevation of bilirubin at one visit and a $\geq 3\text{xULN}$ elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's Law criteria.

Potential hepatotoxicity (meeting one of the PDILI or Hy's Law laboratory criteria at least once) will be considered with and without symptoms potentially associated with hepatitis or hypersensitivity according to the investigator (reported on the Symptoms of Hepatitis and Hypersensitivity CRF page).

Change #74

Section 10.4.1 Vital signs

The following text was added:

Unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

Change #75

Section 10.4.3 Other safety endpoints

The following text was added:

For by-visit summaries, unscheduled and repeat visits will not be summarized, but these data will be included in listings. By-visit tables should include the SFU visit. Summaries over a period of time (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

Change #76

Section 10.4.3.2 Columbia-Suicide Severity Rate Scale (C-SSRS)

The following text was updated:

The incidence of participants with suicidal ideation, suicidal behavior, suicidal ideation or behavior, and self-injurious behavior will be summarized **for the Initial Treatment Period and the combined Initial and Maintenance Treatment Period** by treatment group ~~for each treatment.~~

Change #77

Section 10.3.4.5 Patient Health Questionnaire (PHQ)-9 scores

The following text was updated:

~~A In addition, a categorical summary of the absolute and change from Baseline value scores will be presented by treatment group and visit.~~

~~The percentage of study participants with scores below a corresponding shift table. The following categories will be presented: 0-4; 5, between 5 and 9, between 10 and 14, between 15 and 19, and greater than or equal to 20 in PHQ-9 will be summarized as a shift from Baseline by visit and treatment group based on observed values; ≥ 20 .~~

The percentage of study participants with scores ≥ 15 at any post-Baseline visit and the number and percentage of study participants with scores ≥ 20 at any post-Baseline visit will be summarized by treatment group based on observed values. This summary will also include the percentage of study participants with increase from baseline ≥ 5 at any post-Baseline visit.

~~Different to other safety variables, PHQ-9 will be summarized using the MCMC/monotone regression approach described for continuous variables.~~

The number and percentage of participants that complete the PHQ-9 will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. ~~(or MS, as appropriate).~~ The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit.

Change #78

Section 11 References

The following reference was updated:

Common Terminology Criteria for Adverse Events (CTCAE); Version 4.0 June 2010
2017. U.S. Department of Health and Human Services

Change #79

Section 12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The algorithm for identifying anaphylaxis was updated:

The SMQ *Anaphylactic reaction* consists of three parts:

- A **narrow search** containing PTs that represent core anaphylactic reaction terms (Category A – core anaphylactic reaction terms)

Anaphylactic reaction
Anaphylactic shock
Anaphylactic transfusion reaction
Anaphylactoid reaction
Anaphylactoid shock
Circulatory collapse
Dialysis membrane reaction
Kounis syndrome
Shock
Shock symptom
Type I hypersensitivity

- A **broad search** that contains additional terms that are added to those included in the narrow search. These additional terms are signs and symptoms possibly indicative of anaphylactic reaction and categorized in B, C or D

Category B (Upper Airway/Respiratory Terms)

Acute respiratory failure	Nasal obstruction
Asthma	Oedema mouth
Bronchial oedema	Oropharyngeal spasm
Bronchospasm	Oropharyngeal swelling
Cardio-respiratory distress	Respiratory arrest
Chest discomfort	Respiratory distress
Choking	Respiratory failure
Choking sensation	Reversible airways obstruction

Circumoral oedema	Sensation of foreign body
Cough	Sneezing
Cyanosis	Stridor
Dyspnoea	Swollen tongue
Hyperventilation	Tachypnoea
Irregular breathing	Throat tightness
Laryngeal dyspnoea	Tongue oedema
Laryngeal oedema	Tracheal obstruction
Laryngospasm	Tracheal oedema
Laryngotracheal oedema	Upper airway obstruction
Mouth swelling	Wheezing

▪ **Category C (Angioedema/Urticaria/Pruritus/Flush terms)**

Allergic oedema	Oedema
Angioedema	Periorbital oedema
Erythema	Pruritus
Eye oedema	Pruritus allergic
Eye pruritus	Pruritus generalised
Eye swelling	Rash
Eyelid oedema	Rash erythematous
Face oedema	Rash generalised
Flushing	Rash pruritic
Generalised erythema	Skin swelling
Injection site urticaria	Swelling
Lip oedema	Swelling face
Lip swelling	Urticaria
Nodular rash	Urticaria papular
Ocular hyperaemia	

▪ **Category D (Cardiovascular/Hypotension terms)**

Blood pressure decreased
Blood pressure diastolic decreased
Blood pressure systolic decreased
Cardiac arrest
Cardio-respiratory arrest
Cardiovascular insufficiency

Diastolic hypotension
Hypotension

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after, and for scenarios where two events must have been reported, both events must have occurred within one day of each other:
 - A narrow term or a term from Category A;
 - A term from Category B - (Upper Airway/Respiratory) AND a term from Category C - (Angioedema/Urticaria/Pruritus/Flush);
 - A term from Category D - (Cardiovascular/Hypotension) AND [a term from Category B - (Upper Airway/Respiratory) OR a term from Category C - (Angioedema/Urticaria/Pruritus/Flush)]

Change #80

Section 912.2 Appendix B: Definition of CTCAE grades

Table 12-1 was updated:

Table 12–1: Definitions of CTCAE grades by biochemistry parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine ¹	High	umol/L mmol/L	>1-1.5x Baseline or >ULN-1.5 x ULN	>1.5-3.0x Baseline or >1.5 – 3.0 x ULN	>3.0x Baseline or >3.0 – 6.0 x ULN	>6.0 x ULN
Sodium	Low	mmol/L	130-<LLN	N/A	120-<130	<120
Sodium	High	mmol/L	>ULN-150	>150-155	>155-160	>160
Potassium ²	Low	mmol/L	3.0-<LLN	3.0-<LLN	2.5-<3.0	<2.5
Potassium	High	mmol/L	>ULN-5.5	>5.5-6.0	>6.0-7.0	>7.0
Calcium	Low	mmol/L	2.0-<LLN	1.75-<2.0	1.5-<1.75	<1.5
Calcium	High	mmol/L	>ULN-2.9	>2.9-3.1	>3.1-3.4	>3.4
Magnesium	Low	mmol/L	0.5-<LLN	0.4-<0.5	0.3-<0.4	<0.3
Magnesium	High	mmol/L	>ULN-1.23	N/A	>1.23-3.30	>3.30
Cholesterol	High	mmol/L	>ULN-7.75	>7.75-10.34	>10.34- 12.82	>12.82

1 The CTCAE Grade definitions for creatinine are based on the logical or and the highest applicable CTCAE grade should be assigned to a creatine value.

2 The decreased potassium criterion of 3.0-<LLN is specified for both CTCAE Grade 1 and Grade 2; values meeting this criterion will be counted as Grade 2.

And Table 12-2 was updated:

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin	Low	g/L	100-<LLN	80-<100	<80	N/A
Hemoglobin ¹	High	g/L	>0-20 above ULN or >0-20 above Baseline if Baseline is above ULN	>20-40 above ULN or >20-40 above Baseline if Baseline is above ULN	>40 above ULN or >40 above Baseline if Baseline is above ULN	N/A
Platelets	Low	10 ⁹ /L	75-<LLN	50-<75	25-<50	<25
WBC	Low	10 ⁹ /L	3-<LLN	2-<3	1-<2	<1
WBC	High	10 ⁹ /L	N/A	N/A	>100	N/A
Lymphocytes	Low	10 ⁹ /L	0.8-<LLN	0.5-<0.8	0.2-<0.5	<0.2
Lymphocytes	High	10 ⁹ /L	N/A	>4-20	>20	N/A
Neutrophils	Low	10 ⁹ /L	1.5-<LLN	1.0-<1.5	0.5-<1.0	<0.5

LLN=lower limit of normal; N/A=not applicable; ULN=upper limit of normal, WBC=white blood cells

1 The CTCAE Grade definitions to be applied are dependent on the Baseline hemoglobin value. If the baseline value is > ULN then the criteria relative to Baseline is applicable, otherwise the criteria relative to ULN is applicable.

13.2 Amendment 2

Rationale for the amendment

The main purposes of this amendment were:

- General update to analyses to align with protocol amendment 4.
- Procedural clarifications from discussions and feedback provided at meetings
- Update to align with the bimekizumab program standards and safety topics of interest

Modifications and changes

Global Changes

The following changes were made throughout the SAP:

- Typos and formatting were updated throughout the document
- HSSDD worst pain and average pain were updated to worst skin pain and average skin pain, respectively

Specific changes

In addition to the global changes, the following specific changes have been made (typos such as missing spaces or redundant spaces are not listed):

Change #1

List of Abbreviations

The following abbreviations have been added:

CFB	change from Baseline
CV-CAC	Cardiovascular Event Adjudication Committee
eCDF	empirical cumulative distribution function
IBD-CAC	Inflammatory Bowel Disease Adjudication Committee

Change #2

Section 1 Introduction

The protocols were updated:

The SAP is based on the Protocol Amendment 4 3, 9 May 2022 9 February 2021.

Change #3

Section 2.2 Study endpoints

The following text was deleted:

The endpoints based on HS Symptom Daily Diary (HSSDD) and Hidradenitis Suppurativa Symptom Questionnaire (HSSQ) pain responses are based on the current definitions, which are continuous. It is anticipated that a responder (binary) endpoint will be defined for the HSSDD and HSSQ pain items as well as other symptom items prior to database lock and unblinding, based on separate ongoing, blinded, psychometric analyses aiming to determine threshold for within-patient clinically meaningful improvement.

The below HSSDD and HSSQ pain response endpoints and analyses will be adjusted accordingly in a future SAP amendment.

Change #4

Section 2.2.1.2 Secondary efficacy endpoints

The following bullet was added:

- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline

Change #5

Section 2.2.1.3 Other efficacy endpoints

The following bullets were updated:

- ~~Skin pain response status~~ain response, as assessed by the “worst pain” item in the HSSDD, defined as an improvement from baseline in the weekly worst skin pain score of at least 3 units **Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline**

- **Skin pain response** ~~Response in HS Skin Pain (11-point numeric rating scale) assessed by the HSSDD at Week 16~~ **Response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline**
- **Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline**
- ~~HS Symptom Questionnaire (HSSQ) response (at least a 3-unit reduction from Baseline in worst HS Skin Pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline assessed HSSQ by the HSSDD in the Initial Treatment Period, and assessed by the HSSQ in the Maintenance Treatment Period~~ **Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline**
- Absolute change from Baseline in DLQI Total Score
- DLQI Total Score of 0 or 1
- Minimum clinically important difference (MCID) (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4
- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) and Total score
- Patient Global Impression of HS Severity (PGI-S-HS)
- Patient Global Impression of Change of HS Severity (PGI-C-HS)
- Patient Global Impression of Severity of Skin Pain (PGI-S-SP)
- Patient Global Impression of Change of Skin Pain (PGI-C-SP)
- Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor.
- **Responders Response** on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor

Change #6

Section 2.4 Determination of sample size

The power to detect a statistically significant difference for each of the endpoints are shown in [Table 2-1](#). Notably, with a 2-sided significance level of 0.025, the sample size of 140:70 provides 73% power for detecting at least a difference of 1.5 (bimekizumab Q4W vs placebo) for the Worst **Skin Pain change from Baseline (CFB)** endpoint.

Given the high level of power for each of the primary and key secondary endpoints at the 0.025 significance level for the Q2W comparison (power ≥ 0.89), and per the alpha spending strategy, there is a high likelihood that the Q4W comparison of Worst **Skin Pain CFB** vs placebo ~~for Worst Pain change from Baseline~~ will be allowed to be tested against the 0.05 level of

significance. The power for this latter test is 81%. The sample size is thus ultimately driven by the Worst Skin Pain CFB endpoint. Furthermore, the randomization ratio of 2:2:2:1 has been chosen to provide study participants with a high probability (6/7 ~ 86%) of being randomized to active study drug.

After randomization for this study was complete, an additional endpoint to assess Worst Skin Pain response was included in the sequential testing procedure. This additional endpoint is based on the threshold for clinically meaningful change and is defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline. Note that the power calculations reported in Table 2-1 for this endpoint are based on the sample size that was initially driven by the Worst Pain CFB endpoint as described above. With a 2-sided significance level of 0.025, the sample size of 104:52 in the subset of participants reporting Baseline HSSDD worst skin pain score at or beyond the threshold for clinically meaningful change (ie, Baseline HSSDD ≥ 3) provides 53% power for detecting a statistically significant difference between bimekizumab Q4W and placebo in the proportion of Worst Pain responders.

Note that the power at the 0.025 level of significance associated with this endpoint for the comparison between bimekizumab Q2W and placebo is 95%. The Q4W comparison of Worst Skin Pain response vs placebo against the 0.05 level of significance is therefore likely, and the power at this significance level is 65%. Given the strength of the power for the bimekizumab Q2W arm vs placebo across endpoints, it is considered acceptable to have a relatively low power for this final endpoint in the testing sequence for the bimekizumab Q4W treatment arm vs placebo.

Change #7

Section 2.4 Determination of sample size

Table 2-1 was updated:

Table 2-1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
Flare	0.99	0.99	Proportion of participants with flare by Week 16=0.09	Proportion of participants with flare by Week 16=0.19	Proportion of participants with flare by Week 16=0.52
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6

Table 2–1: Power calculation assumptions and methods

Worst Skin Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0004 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms

^b Within-participant average of Worst Skin Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Skin Pain score at or above 3 (ie, the threshold for clinically meaningful change from Baseline).

Change #8

Section 3.1 General presentation of summaries and analyses

The following text was added:

Per protocol, visit windows are ± 3 days from the date of first dose. The 20-week SFU Visit window is ± 7 days from the date of the final dose. All by-visit summaries will contain nominal (ie, scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for assessments that occur within a 3-day time window of a scheduled visit. In that case, the assessment will be mapped to the corresponding scheduled visit and will be used for the analysis. This will only occur for selected vendor data. **The only exception to this rule is for unscheduled assessments that occur up to 3 days after the Baseline visit. These unscheduled visits will remain as unscheduled as the Baseline assessment cannot be after the first dose of study drug administration. See Section 3.3 for more details on the definition of Baseline values.**

Change #9

Section 3.5.8 Pharmacokinetics Per-Protocol Set

The following text was deleted:

The Pharmacokinetics Per-Protocol Set (PK-PPS) will consist of study participants who received at least 1 full dose of bimekizumab and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the PK. ~~The Pharmacokinetics Per-Protocol Set is defined separately for each of the treatment periods (ie, separately for the Initial Treatment Period and the Maintenance Treatment Period).~~

Change #10

Section 3.10 Changes to protocol-defined analyses

The following text was updated:

The following other efficacy endpoints are included in the protocol but will not be included as part of the analysis:

- **Responders-Response** on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48

The calculation of nominal p-values has been added for selected efficacy endpoints. **These nominal p-values are not controlled for multiplicity and should not be used to declare statistical significance.**

The protocol defines the PK-PPS separately by period, but there will only be one PK-PPS for the overall study.

Change #11

Section 4.1 Adjustments for covariates

The following sentence was updated:

The Worst Skin Pain **secondary endpoints (change from Baseline continuous secondary endpoint and pain response binary endpoint)** will also include analgesic use as a covariate.

Change #12

Section 4.2.1.2 Handling of missing data for the secondary efficacy endpoints

The following paragraph was updated:

For secondary continuous efficacy endpoints, MI-MCMC/monotone regression is the primary method for imputing missing data, regardless of whether the missing data are preceded by an intercurrent event. That is, if an intercurrent event occurs on or before a visit, the result for that visit will be treated as missing and **then imputed with the missing data**. If the imputation model cannot converge, last observation carried forward (LOCF) will be used. The OC method will be performed as a sensitivity analysis.

Change #13

Section 42.2.1 MI – MCMC/Monotone Regression

The imputation rule for HSSDD was updated in Table 4-2.

Table 13–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
Lesion count ^a	0	--	Yes
DLQI total score	0	30	Yes
hs-CRP	LLOQ/2	--	No
HSSDD item score	0	10	NoYes
HSSQ item score	0	10	Yes
HiSQOL symptom status score	0	16	Yes

Table 13–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
HiSQOL psychosocial impact score	0	20	Yes
HiSQOL impact on physical activities score	0	32	Yes
EQ-5D-3L VAS	0	100	Yes
WPAI dimension scores	0	100	No for variables: “Percent work time missed due to problem” and “Percent overall work impairment due to problem” Yes for variables: “Percent impairment while working due to problem” and “Percent activity impairment due to problem”. These two variables can only take values that are multiples of 10.

^a Lesion counts will be imputed separately for each lesion type (abscesses, draining tunnels [fistulas/sinus tracts], inflammatory nodules, non-draining tunnels [fistulas/sinus tracts], non-inflammatory nodules, HS scars). The imputed lesion counts will be used to derive the endpoints that are dependent on the lesion count data (eg, HiSCR₅₀).

Change #14

Section 4.2.3 Rationale for estimand

The following text was added to the bullet:

- A composite estimand strategy will be used for the primary analysis of the primary and binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, flare, **HS worst skin pain response**),

Change #15

Section 4.5 Multiple comparisons/multiplicity

Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$).

Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

1. Step 1: Test HiSCR₅₀ at significance level 0.025.
2. Steps 2 to ~~65~~ – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown in [Figure 4-1](#), moving to the next step only if significance achieved at 0.025.
3. In the event that Step ~~65~~ is significant at 0.025 for a given dose, then Steps 1 to 6 will be repeated for the other dose using a significance level of 0.05.

The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment using the analysis methods specified in Section 8.3:

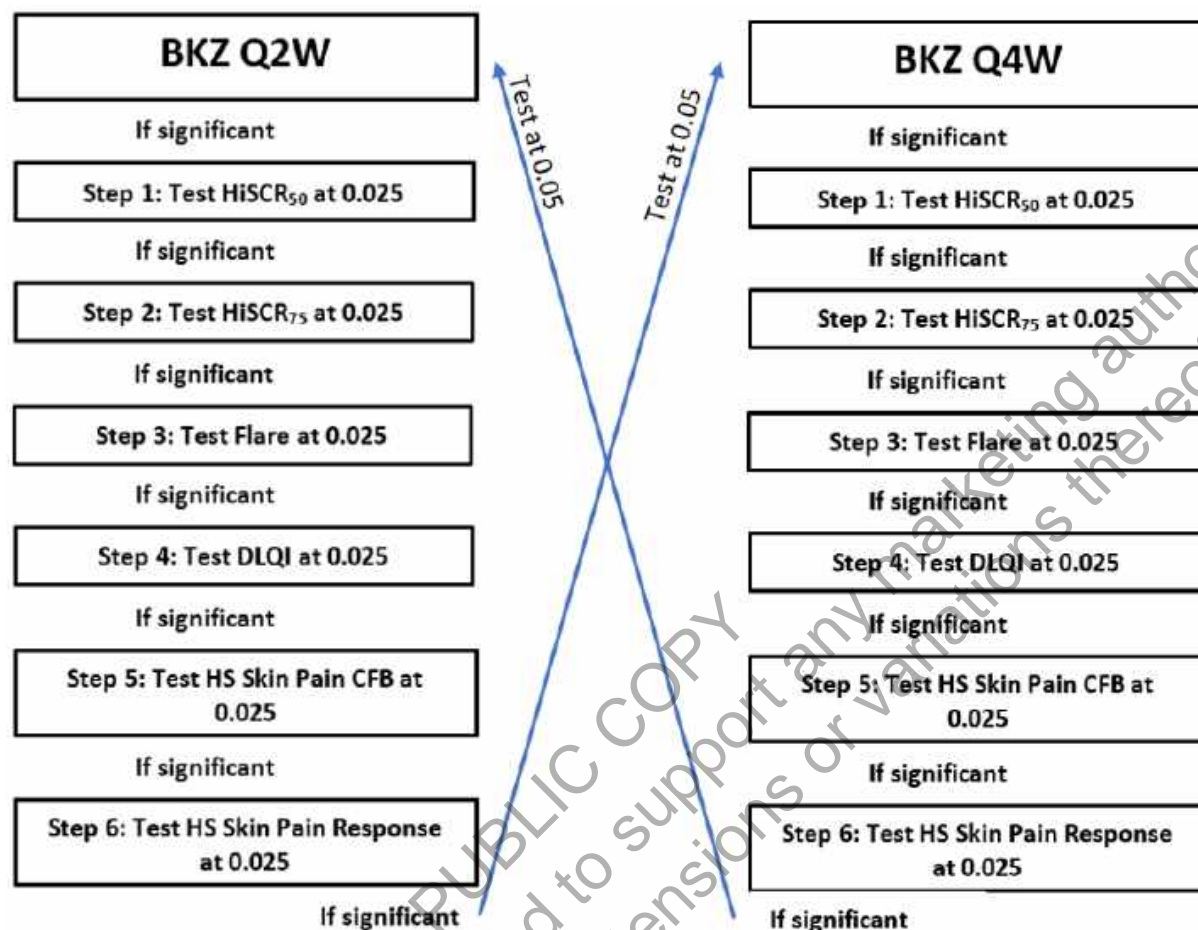
1. Proportion of study participants who achieve HiSCR₇₅ at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
2. Proportion of study participants who experience at least 1 flare by Week 16, with flare defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
3. CFB in DLQI Total Score at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
4. Absolute change from Baseline in Skin Pain Score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HSSDD.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
5. **Skin pain response at Week 16, based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline.**
 - a. **bimekizumab 320mg Q2W vs placebo**
 - b. **bimekizumab 320mg Q4W vs placebo**

Change #16

Section 4.5 Multiple comparisons/multiplicity

Figure 4-1 was updated to add the new secondary endpoint:

Figure 4-1: Sequence of testing



AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks

HS skin pain response is tested among study participants with a score of ≥ 3 at Baseline.

Change #17

Section 4.8 Examination of subgroups

The following text was updated:

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, ~~and flare~~, **and worst skin pain response** endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period. Additional subgroup analyses will be performed on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, ~~and flare~~, **and skin pain response endpoints** which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

Additionally, the following bullets were clarified:

- Antibody positivity (confirmatory assay: negative or positive. **See Section 9.3.2**)
- **Antihistamines users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users) (applicable only to the skin pain response endpoint)**

Change #18

Section 6.4 Prior and concomitant medications

The following text was added:

The number and percentage of study participants with concomitant vaccines for COVID-19 will be summarized by treatment group, overall and by World Health Organization Drug Dictionary Standardized Drug Grouping (SDG), presenting SDG subgroup, and preferred term. The SDG subgroup Vaccines for COVID-19 will be used to identify vaccines for COVID-19 using the narrow scope; this subgroup is divided further into separate subgroups which is the level that will be presented. The number of individual occurrences of the vaccine for COVID-19 will also be summarized.

A listing of concomitant vaccines for COVID-19 will be provided.

Change #19

Section 6.4.2 Classification of participants as analgesic, antihistamine users

The section was updated as follows:

If a participant has taken a new analgesic/increased regimen of analgesic, or taken an antihistamine, on 1 or more days (need not be consecutive) ~~for a given week in a study period (Initial Treatment Period or Maintenance Treatment Period), then for that week period the participant will be classified as an analgesic or antihistamine user, respectively. The week period under consideration is to match the period as defined for the HSSDD week for the Initial Treatment Period or HSSQ week for the Maintenance Treatment Period, based on dates/times of the medications taken. If there is a visit date but no HSSDD available at the visit, then the analgesic/antihistamine user status for that week will be derived based on the visit date. If there is no visit available, then the weekly analgesic/antihistamine user status will default to the analgesic/antihistamine status for the overall study period.~~

New analgesic/increased regimen of analgesic use, regardless of indication, is defined as an analgesic medication with start date on or after the first dose of study medication. Stable analgesics (ie, analgesics which were taken already before randomization) will not be included in this category of analgesic user. This classification will be used ~~to adjust the formal analysis of the Worst Pain secondary endpoints and for selected subgroup analyses.~~

Antihistamine use is identified by considering the ATC classification. This classification is used for analyzing the Worst Itch endpoint and for selected subgroup analyses, by visit, for the Initial Treatment Period and Maintenance Period as applicable.

Additionally, if a participant has taken a new analgesic/increased regimen of analgesic on 1 or more days (need not be consecutive) prior to the Week 16 visit, then for that week the participant will be classified as an analgesic user. This classification will be used to adjust the formal analysis of the Worst Skin Pain secondary endpoints. If there is a visit date but

no HSSDD available at the visit, then the analgesic/antihistamine user status for that week will be derived based on the visit date. If there is no visit available, then the weekly analgesic/antihistamine user status will default to the analgesic/antihistamine status for the overall study period.

Change #20

Section 8.2 Primary efficacy endpoint

The intercurrent event strategy for the sensitivity analysis was updated in Table 8-1:

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Statistical Category (Section)	Estimands for Primary Endpoint			
	Variable/Endpoint	Pop	IES	PLS (Analysis)
Primary Objective: To evaluate the efficacy of bimekizumab in study participants with moderate to severe HS				
Sensitivity (Section 8.2.3.2)	HiSCR ₅₀ response at Week 16	RS	Composite strategy, as for the primary analysis. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – Reference-Based Regression under a missing not at random assumption.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Statistical Category (Section)	Estimands for Primary Endpoint			
	Variable/Endpoint	Pop	IES	PLS (Analysis)
Sensitivity (Section 8.2.3.5)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

Change #21

Section 8.2.3.5 Analysis on observed cases

The following text was updated:

An additional supportive analysis will be based on observed data only for study participants who are still on the randomized treatment at Week 16. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing (see Section 4.2.2). **participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data at Week 16 will be treated as missing (see Section 4.2.2).**

Change #22

Section 8.2.3.11 Center-by-Treatment Interaction

The following sentence was added: In order to achieve model convergence, other explanatory variables eg, Baseline Hurley Stage and Baseline antibiotic use may be dropped from the model.

If model convergence is still not achieved, region and a region-by-treatment interaction term will be added to the model instead. Regions are defined in Section 3.7.

Change #23

Section 8.3 Secondary efficacy endpoints

The following sensitivity analyses as well as the new skin pain response secondary analysis were added to Table 8-2:

Table 8–2: Estimand Details and Attributes for Secondary Analyses

		Estimands for Secondary Endpoints			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Secondary Objective: Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS					
DLQI	Secondary - Sensitivity (Section 8.3.2.2)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a DLQI total score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.3.2)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.
HSSDD	Secondary (Section 8.3.4.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary Sensitivity (Section 8.3.4.2.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.
HSSDD	Secondary Sensitivity (Section 8.3.4.2.2)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

AE=adverse event; ANCOVA=analysis of covariance; DLQI=Dermatology Life Quality Index; HiSCR=Hidradenitis Suppurativa Clinical Response; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; IES=intercurrent event(s) strategy; LSMD=Least Squares Mean Difference; MCMC=Markov Chain Monte Carlo; MI=multiple imputation; PLS=Population-level summary; Pop=Population; RS=Randomized Set

^a **Analysis includes all study participants in the RS with a Baseline HSSDD Worst Skin Pain score of 3 or higher.**

Change #24

Section 8.3.3.1 Primary analysis of change from Baseline in DLQI Total Score

Section 8.3.3.1 was separated from Section 8.3.3 due to the added sensitivity analysis for DLQI Total Score. The following text was added:

Change from Baseline in DLQI total score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use and Baseline value as a covariate. The least square mean (LSM), standard error (SE), and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab: the difference between the LSM, the associated 97.5% CI for the contrasts, and the corresponding p-value will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96. **Estimand and intercurrent event details are specified in Table 8–2.**

Change #25

The following section was added:

Sensitivity analysis of change from Baseline in DLQI Total score at Week 16 (Section 8.3.3.2)

A sensitivity analysis using the same analysis model as in Section 8.3.2.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

Change #26

Section 8.3.4.1 Primary analysis of change from Baseline in skin pain score at Week 16

Section 8.3.4.1 was separated from Section 8.4.3 due to the added sensitivity analysis for HSSDD worst skin pain score and removed ‘A treatment by analgesic use interaction term will also be added to the model and removed if not significant’ from section.

Change #27

The following section was added:

Sensitivity analysis of change from Baseline in skin pain score at Week 16 (Section 8.3.4.2)

A sensitivity analysis using the same analysis model as in Section 8.3.3.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

Change #28

The following sections were added:

HSSDD skin pain response at Week 16 (Section 8.3.5)

The analysis set for the analyses of the skin pain response will be restricted to those study participants in the RS with a Baseline worst skin pain score of 3 or higher. The weekly scores and Baseline score are derived as specified in Section 8.3.3.

Primary analysis of skin pain response at Week 16 (Section 8.3.5.1)

Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, is defined as an improvement in the weekly worst skin pain score of at least 3 points versus Baseline.

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, Baseline antibiotic use, and analgesic use (Section 6.4.2).

The odds ratio versus placebo, p-value (from Wald test), and 97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose. Missing data will be handled as specified in Section 4.2.1.2. Estimand and intercurrent event details are specified in Table 8–2.

The number and percentage of participants who are pain responders at Week 16 will be summarized by treatment group.

By-participant listings of pain response status will be provided.

Sensitivity analyses of Skin Pain Response at Week 16 (Section 8.3.5.2)

Nonresponse imputation (Section 8.3.5.2.1)

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (Table 8–2) will be imputed as nonresponse. That is, participants who experience an intercurrent event will be imputed as nonresponder at the timepoint of the event and all subsequent timepoints (including any recorded data after the event), and all missing data will also be imputed as nonresponse.

The same analysis model as Section 8.3.4.1 will then be used on the imputed data set.

Analysis on observed case (Section 8.3.5.2.2)

An additional supportive analysis will be based on observed data only for study participants with a worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing (see Section 4.2.2).

The same analysis model as in Section 8.3.4.1 will then be used on the imputed data set.

Change #29

Section 8.4.1.3 HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response at both Weeks 16 and 48

The following text was added:

The number and percentage of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at both Weeks 16 and 48 will be summarized based on the RS and MS.

Change #30

Section 8.4.12 Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)

The following text was updated:

See Section 8.3.3 for details on HSSDD Baseline and weekly average definitions and derivations.

Percent change from Baseline in HSSDD responses for worst and average skin pain score is defined as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline HSSDD score} - \text{Baseline HSSDD score}}{\text{Baseline HSSDD score}}$$

Change from Baseline in each HSSDD item (worst skin pain, average skin pain, smell or odor, itch at its worst, and amount of drainage or oozing) score will be summarized using descriptive statistics by treatment group and visit, based on weekly averages. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits. Percentage change will be summarized for the worst and average skin pain items.

Additionally, change from Baseline in each HSSDD item will be evaluated by treatment group at Week 16 via continuous empirical cumulative distribution function (eCDF) plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Change from Baseline in Worst **Skin** Pain score and Worst Itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

HSSDD response **based on clinically meaningful change** for the worst skin pain and average skin pain items is defined as at least a **3-point** 30% reduction and at least a 1-point reduction from Baseline in HSSDD among study participants with a score of ≥ 3 at Baseline, based on weekly averages.

The number and percentage of responders based on clinically meaningful change for **each item the worst skin pain item** will be summarized by treatment group and visit.

The number and percentage of participants who were responders **based on clinically meaningful change** at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst ~~HSSDD~~ skin pain score **item**.

HSSDD response for the worst skin pain and average skin pain items is defined as at least a 30% reduction and at least a 1-point reduction from Baseline among study participants with a score of ≥ 3 at Baseline. The number and percentage of responders for each item will be summarized by treatment group and visit.

The number and percentage of participants who were responders based on clinically meaningful change at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst skin pain and average skin pain items.

The number and percentage of participants that complete the HSSDD will be calculated for each visit by treatment group. A participant will be counted as completing the HSSDD at a visit if the minimum number of daily entries is present to calculate the weekly average (see Section 8.3.3). The percentage will be based on the number of participants in the RS. A participant will be considered a completer at a visit if the weekly average can be calculated for that visit.

Change #31

Section 8.4.13 Hidradenitis Suppurativa Symptom Questionnaire

The following text was updated:

Additionally, change from Baseline in each HSSQ item will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

HSSQ response for skin pain item is defined as at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline.

Change #32

Section 8.4.15 Hidradenitis Suppurativa Quality of Life

The following text was added:

Additionally, change from Baseline in each HiSQOL subscale will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Change #33

Section 8.5 Additional statistical analyses of other efficacy endpoints

The following analysis was added

- Skin Pain response per HSSDD at Week 12

Change #34

Section 9.3.2 Anti-bimekizumab antibodies

The section was updated as follows:

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-tiered assay approach: screening assay, confirmatory assay, and titration assay. Samples confirmed as positive within the confirmatory assay will be further evaluated for the presence of neutralizing anti-bimekizumab antibodies specific to IL-17AA, IL-17FF or both. Samples will be taken at Baseline, then at study Weeks 4, 8, 12, 16, 20, 24, 36 and 48, and at PEOT and SFU timepoints.

ADAb samples are not analyzed when study participants are on a treatment other than bimekizumab. For study participants who switch from placebo to bimekizumab, samples are analyzed starting at the visit when the switch to bimekizumab occurs (Week 16). The sample at Week 16 will act as the Baseline for that treatment group.

The screening cut point will be used to determine the status of anti-bimekizumab antibodies in the test sample as Positive Screen (PS) or Negative Screen (NS). For samples presenting anti-bimekizumab antibody levels that are PS, a further confirmatory assay will be performed, and the result of which will be reported as either Positive Immunodepletion (PI) or Negative Immunodepletion (NI).

ADAb status for each sample will be derived as follows:

- Sample values that are either NS, or PS and NI and where the bimekizumab concentration is less than the validated ADAb assay drug tolerance limit will be defined as anti-bimekizumab antibody negative.
- Sample values that are either NS, or PS and NI and where the bimekizumab concentration exceeds the validated ADAb assay drug tolerance limit will be defined as inconclusive.
- Sample values that are PS and PI will be defined as ADAb positive (regardless of availability of a titer value)
- Missing or non-evaluable samples will be defined as missing

Positive immunodepletion samples will be titrated, and the ADAb titer (reciprocal dilution factor including minimum required dilution) will be reported. Subsequently, PI samples will also be subject to a neutralizing assay to evaluate the potential of ADAb to neutralize the target binding of bimekizumab (IL-17AA or IL-17FF or both) in vitro.

Cumulative ~~There are different levels of classification for ADAb status~~ **will be derived as follows:**

The ADAb status (positive, negative or missing) will be considered in a cumulative manner at each time point.

A study participant will be counted positive from the first visit at which the study participant achieved a positive ADAb sample result to the end of the treatment period, regardless of any missing/inconclusive or negative ADAb sample result.

If a study participant has only negative ADAb samples or only one missing/inconclusive sample with all other ADAb samples being negative, the study participant will be classified as negative. An exception remains for the Baseline Visit where only one sample could be available. If the sample is missing/inconclusive, then the sample will be classified as being negative for the cumulative ADAb status.

Otherwise, the study participant will be classified in the missing ADAb category.

Overall ADAb status will be derived as follows:

A study participant will be classified as:

- **Positive if the study participant has at least one positive sample up to the time point of interest (regardless of having missing/inconclusive data).**
- **Negative if the study participant has all the samples negative or only one missing/inconclusive sample with negative ADAb samples up to the timepoint of interest.**
- **Missing if the study participant has more than one missing ADAb result (or have more than one inconclusive sample) and all other available ADAb samples are negative up to the time point of interest.**

ADAb categories will be derived definitions are as follows:

- **Pre ADAb negative – treatment-emergent ADAb negative (Category 1):** Includes study participants who are anti-bimekizumab antibody negative at Baseline and anti-bimekizumab antibody negative at all sampling points during the period of interest (one post-Baseline

missing/inconclusive sample is allowed for subjects with pre- anti-bimekizumab antibody negative sample). This group also includes study participants who have a missing or inconclusive sample (either missing or inconclusive or insufficient volume) at Baseline (ie, pre-treatment) with all post-Baseline samples as ADAb negative.

- **Pre ADAb negative – treatment-emergent ADAb positive (Category 2):** Includes study participants who are ADAb negative at Baseline and ADAb positive at any sampling points post-Baseline during the period of interest. This group also includes study participants who have a missing sample (either missing or insufficient volume) at Baseline (ie, pre-treatment) with 1 or more post-Baseline samples as ADAb positive.
- **Pre ADAb positive – treatment-emergent reduced ADAb (Category 3):** Includes study participants who are ADAb positive at Baseline, and ADAb negative at all sampling points post-Baseline during the period of interest.
- **Pre ADAb positive – treatment-emergent unaffected ADAb positive (Category 4):** Includes study participants who are ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with titer values of the same magnitude as Baseline (ie, less than a predefined fold difference from the Baseline titer).
 - For this analysis, this is set at an increase of less than the validated Minimum Significant Ratio (MSR) of 2.07-fold from Baseline.
- **Pre ADAb positive – treatment-emergent ADAb boosted positive (Category 5):** Includes study participants who ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with increased titer values compared to Baseline (equal to or greater than a predefined fold difference increase from Baseline titer which will be defined within the validation of the assay).
 - For this analysis, this is set at an increase equal to or greater than the validated MSR of 2.07-fold from Baseline.
 - Note: for any study participant who is ADAb positive at Baseline and ADAb positive at a post-Baseline time point during the period of interest, but for whom titers are not available to determine treatment unaffected or treatment boosted status, the study participant will be considered as treatment boosted, assuming no other samples are available.
- **ADAb Inconclusive (Category 6):** Includes study participants who have an ADAb positive Baseline (pre-treatment) sample and some post-Baseline samples during the period of interest are missing or inconclusive, while other post-Baseline samples are ADAb negative.
- **Total treatment-emergent ADAb positivity (Category 7 [Categories 2 and 5 combined]):** Includes study participants who are pre ADAb negative – treatment-emergent ADAb positive (Category 2) and pre ADAb positive – treatment boosted ADAb positive (Category 5).
- **Total prevalence of pre- ADAb positivity (Category 8 [Categories 3, 4, 5 and 6 combined]):** Study participants that are tested ADAb positive at Baseline.
- **Missing:** Includes study participants who are ADAb negative, missing, or inconclusive at Baseline with some post-Baseline samples as missing or inconclusive, and other samples as ADAb negative.

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, the following categories can also be used:

- **ADAb positive – This is defined as study participants who are anti-bimekizumab antibody positive on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).**
- **ADAb negative – Study participants for who either:**
 - **All samples (including Baseline) are ADAb negative and there are no missing or inconclusive samples**
 - **Only 1 sample is ADAb positive and all other samples (including Baseline) are ADAb negative or missing or inconclusive**
 - **Only 1 sample is missing or inconclusive and the remaining ADAb samples are negative.**
- **ADAb missing - Defined as study participants who do not fulfil the criteria for one of the 2 groups listed above.**

The rationale for requiring at least 2 time points in which ADAb levels are above the specified cut point is to exclude those study participants who have only one occurrence of ADAb levels during the course of treatment. Including such study participants would increase the number of ADAb positive study participants with potentially no impact on efficacy.

In the case that a sample is collected 1 or more days following the scheduled visit date in which the drug was administered, the ADAb results for that sample will be associated with the scheduled visit and summarized accordingly. Such samples will also be considered when ADAb results are summarized over a given study period.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by participants. All analyses will be run on the AMS, unless specified otherwise.

- Summary of ADAb status overall and by each visit separated by treatment group
- Summary of the time-point of the first occurrence of ADAb positivity during the treatment period by treatment group. A plot of the titer by time to first ADAb positivity will be prepared.
- All individual participant-level ADAb results will be listed.
- The number and percentage of participants in each of the 8 ADAb categories during the treatment period by treatment group.
- The prevalence of immunogenicity, separated by treatment group, and defined subcategory, will be reported by visit, defined as (cumulative) proportion of participants having confirmed positive ADAb samples at any visit up to and including that visit. Missing samples will not be included in the denominator.

- The time to achieving treatment-emergent ADA_b positivity, separated by treatment group and defined subcategory, will be analyzed based on Kaplan-Meier methods. **This will be shown only for Categories 2 and 8 above.** Participants will be considered to have an event at the time point at which treatment emergent ADA_b positive is first achieved (taking the MSR into consideration for sub-category 5). Participants classified as treatment-emergent ADA_b negative will be censored at the time of the last available ADA_b result.
- A summary of HiSCR₅₀ responders at Week 16, separated by treatment group, as a function of ADA_b titer will be presented graphically. ~~This will be repeated for HiSCR₇₅ responders.~~
- Individual plots of plasma bimekizumab concentrations/ ADA_b titer both plotted on the Y-axes by visit (x-axis) for the full treatment period (excluding SFU for interim analyses and including SFU for final analyses) will be presented for participants with and without HiSCR₅₀ response at Week 16.
- Spaghetti plots of ADA_b titer (y-axis) by visit (x-axis), separated by treatment group for all ADA_b positive participants, including Baseline positive participants.
- Box plots of ADA_b titer (logscale) by time to first ADA_b positivity by treatment group.

~~For purposes of efficacy subgroup analyses based on anti bimekizumab antibody status, 2 categories will be used:~~

- ~~• ADA_b positive – This is defined as participants who have ADA_b levels above the specified cut point on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).~~
- ~~• ADA_b negative – Participants who are not defined as anti bimekizumab positive (as described above) will be defined as ADA_b negative.~~

The groups for defining ADA_b status for safety subgroup analyses are as follows:

- AEs starting before first ADA_b positive result
- AEs starting on or after first ADA_b positive result
- AEs for participants who were always ADA_b negative

This is further explained in Section 10.2.2.

Change #35

The following section was added:

COVID-19 related considerations (Section 10.2.1.1)

To assess the impact of COVID-19 mass vaccination on TEAEs, a sensitivity analysis will present all TEAEs excluding TEAEs assessed as exclusively related to COVID-19 vaccine by the investigator. TEAEs recorded as related to both study medication and COVID-19 vaccination should not be excluded. A complementary table and listing of TEAEs related to COVID-19 vaccine will be presented.

Another sensitivity analysis will present all TEAEs excluding TEAEs with start date on or up to 5 days after date of COVID-19 vaccine. Note that study participants may receive more than one

administration of COVID-19 vaccine. A complementary table and listing of TEAEs with start date on or up to 5 days after date of COVID-19 vaccine will also be presented.

Change #36

Section 10.2.2 AE summaries

The following AE summaries were added:

- Incidence of TEAEs Excluding TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval Censored TEAEs by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval TEAEs by SOC, HLT, and PT

Change #37

Section 10.2.2 AE summaries

The following text was added:

The following table will be presented for the combined Initial and Maintenance Treatment Period. **This summary will include only AEs that occur while a participant is on bimekizumab. Any AEs in the Initial Treatment Period that begin while a participant is on placebo will be excluded.**

- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status. **This will include columns for the following:**
 - **TEAEs starting before the first ADAb positive result (includes ADAb categories 2 and 5) where TEAEs have occurred before the following events: a) the first positive ADAb result for subjects in category 2 and b) the first post-Baseline boosted ADAb titer result for subjects with titer results and the first post-Baseline positive ADAb result for subjects with positive ADAb at Baseline with no other samples with titer available for subjects in category 5**
 - **TEAEs starting on the same date or after the first ADAb positive result (includes ADAb Categories 2, 3, 4, 5 and 6) where TEAEs have occurred on or after the following events: a) the first positive ADAb results for subjects in categories 2, 3, 4 and 6, and b) the first post-Baseline boosted ADAb titer result for subjects with titer results and the first post-Baseline positive ADAb result for subjects with positive ADAb at Baseline with no other samples with titer available for subjects in category 5**
 - **TEAEs for subjects who are ADAb negative at all timepoints (includes ADAb Category 1)**

The tables with risk differences will also be accompanied by figures (dot plots) which show the incidence of the adverse events and corresponding 95% risk difference confidence intervals. These will be ordered by descending order of risk difference (bimekizumab vs placebo).

Change #38

Section 10.2.3.3 Major adverse cardiac event

The entire section was updated:

Potential cardiovascular events are adjudicated by the independent Cardiovascular Event Adjudication Committee (CV-CAC) according to the CV-CAC Charter (version 6.0). Adjudicated events are classified by the CV-CAC to one of the event types as defined in [Table 10–1](#). The classification of an event as a Major Adverse Cardiac Event (MACE) is also determined by the CV-CAC. Events which are classified by the CV-CAC as any of the event types identified in the third column of [Table 10–1](#) will be considered an extended MACE. Note that extended MACE is determined programmatically and includes a broader scope definition of MACE.

MACE as determined by the CV-CAC will be presented in a table and listing. Extended MACE will be presented separately in a table and listing.

Another table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be a MACE and/or an extended MACE.

Additionally, a listing of all events identified for potential review by the CV-CAC will be produced. This listing will indicate whether each event was identified by the CV-CAC Chair for full committee review.

Table 13–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
1	Non-Fatal Myocardial Infarction (MI)	Yes
2	Non-Fatal Stroke: hemorrhagic	Yes
3	Non-Fatal Stroke: ischemic	Yes
4	Non-Fatal Stroke: embolic	Yes
5	Non-Fatal Stroke: undeterminable	Yes
6	Hospitalization or ER for Unstable Angina with urgent revascularization	Yes
7	Hospitalization or ER for Unstable Angina without urgent revascularization	No
8	Hospitalization for Heart Failure	Yes
9	Transient Ischemic Attack (TIA)	No
10	Coronary Revascularization Procedures (e.g. percutaneous coronary intervention, coronary artery bypass grafting)	Yes

Table 13–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
11	Urgent Revascularization Procedures (i.e. due to symptoms of brain ischemia or pending infarction)	Yes
12	Arrhythmia (not associated with ischemia)	No
13	Peripheral Arterial Event	No
14	Venous Thromboembolic Event: pulmonary embolism (PE)	No
15	Venous Thromboembolic Event: deep vein thrombosis (DVT)	No
16	Venous Thromboembolic Event: PE and DVT	No
17	Other CV Event	No
18	Death due to Myocardial Infarction (MI)	Yes
19	Death due to Stroke	Yes
20	Sudden Cardiac Death	Yes
21	Other CV Death (e.g. heart failure, pulmonary embolism, cardiovascular procedure-related)	Yes
22	Cardiovascular: Undetermined Cause of Death (i.e. cause of death unknown)	Yes
23	Non-Cardiovascular Death	No
24	Non-Cardiovascular Event	No
99	Inadequate information to adjudicate	No

CV=Cardiovascular; DVT=Deep Vein Thrombosis; ER=Emergency Room; MACE= Major Adverse Cardiac Event; MI=Myocardial Infarction; PE=Pulmonary Embolism; TIA=Transient Ischemic Attack.

MACE is determined by the adjudication committee and is not identified programmatically based on event type.

Change #39

Section 10.2.3.5 Suicidal Ideation and Behavior

The entire section was updated:

Potential neuropsychiatric events are adjudicated by the independent Neuropsychiatric Adjudication Committee according to the Neuropsychiatric Adjudication Committee (version 8.0). Adjudicated events are classified by the Committee as Suicidal or Non-suicidal.

Adjudicated events are also further classified by the Committee to one of the event types as

defined in [Table 10–2](#). Suicidal Ideation and Behavior (SIB) is defined as events classified by the Committee as Suicidal.

A table and listing will present SIB events.

Another table and listing will present the adjudicated neuropsychiatric events by type. For each neuropsychiatric event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be Suicidal or Non-Suicidal. For event type suicidal ideation, the listing will also indicate if intent was present and if the suicidal ideation was clinically significant.

Additionally, a listing of all events identified for potential review by the Committee will be produced. This listing will indicate whether each event was identified by the Neuropsychiatric Event Adjudication Committee Chair for full committee review.

Table 13–2: Neuropsychiatric event classifications

Event Type Code	Event Classification	Event Type
1	Suicidal	Suicidal events/completed suicide
2	Suicidal	Suicide attempt
3	Suicidal	Preparatory acts toward imminent suicidal behavior
4	Suicidal/Non-suicidal ^a	Suicidal ideation
7	Non-suicidal	Nonsuicidal Self-injurious behavior
8	Non-suicidal	Nonsuicidal Other
99	Not applicable	Inadequate information to adjudicate

^a Suicidal ideation event types can be classified by the Neuropsychiatric Adjudication Committee as Suicidal or Non-suicidal depending on whether intent to die was present.

Change #40

Section 10.2.3.6 Inflammatory bowel disease

The entire section was updated:

Selected gastrointestinal events are adjudicated by the independent Inflammatory Bowel Disease (IBD) Adjudication Committee (IBD-CAC) according to the IBD-CAC Charter (version 3.0). Adjudicated events are classified by the IBD-CAC into one of the diagnostic types as defined in [Table 10–3](#). The events will further be classified as definite, probable or possible IBD.

An overview of adjudicated IBD events will be stratified by subjects with or without a previous medical history of IBD. Previous medical history of IBD will be determined using the information recorded on the History of IBD CRF page ("Does subject have a history of IBD?"). This overview table will present events adjudicated by the IBD-CAC as either possible, probable or definite IBD. Definite and probable IBD will also be aggregated and summarized in this table.

In addition, this table will summarize each IBD event classification (possible, probable or definite) separately.

Another table and listing will present the adjudicated IBD events by type. For each IBD event type, the individual PTs which fall within each event type will be summarized.

Additionally, a listing of all events identified for potential review by the IBD-CAC will be produced. This listing will indicate whether each event was identified by the IBD-CAC Chair for full committee review.

A further supportive listing will present the individual diagnostic criteria met for each adjudicated IBD event.

Table 13–3: IBD event classifications

Event Type Code	Event Type (Classification and diagnosis)	Classification
1	Possible Inflammatory Bowel Disease – Crohn’s Disease	Possible
2	Probable Inflammatory Bowel Disease – Crohn’s Disease	Probable
3	Definite Inflammatory Bowel Disease – Crohn’s Disease	Definite
4	Possible Inflammatory Bowel Disease – Ulcerative Colitis	Possible
5	Probable Inflammatory Bowel Disease – Ulcerative Colitis	Probable
6	Definite Inflammatory Bowel Disease – Ulcerative Colitis	Definite
7	Possible Inflammatory Bowel Disease – type unclassified	Possible
8	Probable Inflammatory Bowel Disease – type unclassified	Probable
9	Definite Inflammatory Bowel Disease – type unclassified	Definite
10	Symptoms not consistent with Inflammatory Bowel Disease	Not applicable
11	Possible Inflammatory Bowel Disease – Microscopic Colitis	Possible
12	Probable Inflammatory Bowel Disease – Microscopic Colitis	Probable
13	Definite Inflammatory Bowel Disease – Microscopic Colitis	Definite
14	Possible Inflammatory Bowel Disease – no further differentiation possible	Possible
15	Probable Inflammatory Bowel Disease – no further differentiation possible	Probable
16	Definite Inflammatory Bowel Disease – no further differentiation possible	Definite
99	Not enough information to adjudicate	Not applicable

IBD=inflammatory bowel disease.

Note: IBD diagnoses of “microscopic colitis” and “no further differentiation possible” were added in an adjudication charter amendment, accounting for the event type numbering.

Change #41

Section 10.2.3.8 Hepatic events and PDILI

The following word was added:

Cases of **potential** Hy's Law will be reported separately in a liver function test table.

Change #42

Section 12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The following text was added:

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after, and for scenarios where two events must have been reported, both events must have occurred within one day of each other (**as anaphylaxis is an acute event, imputed dates should not be used in the algorithmic approach**):

13.3 Amendment 3

Rationale for the amendment

The main purpose of this amendment was:

- Change the Flare by Week 16 endpoint from a secondary endpoint included in the statistical hierarchy to an other efficacy endpoint, in alignment with changes described in protocol amendment 5
- Clarify the tipping point analysis procedure

Modifications and changes

Global Changes

The following changes were made throughout the SAP:

- Moving the Flare by Week 16 endpoint from any description of secondary efficacy endpoints, and re-inserting this endpoint as an other efficacy endpoint

Specific Changes

In addition to the global changes, the following specific changes have been made (typos such as missing spaces or redundant spaces are not listed):

Change #1

Section 1 Introduction

The protocols were updated:

The SAP is based on the Protocol Amendment 5, 27 September 2022.

Change #2

Flare by Week 16 was removed as a secondary efficacy endpoint:

2.2.1.2 Secondary efficacy endpoints

The secondary efficacy endpoints are defined as:

- HiSCR₇₅ response (defined as at least a 75% reduction from Baseline in the total AN count with no increase from Baseline in abscess or draining tunnel count) at Week 16
- ~~Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16~~
- Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16
- Absolute change from Baseline (CFB) in Skin Pain score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HS Symptom Daily Diary (HSSDD)
- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline

Change #3

Flare by Week 16 was added as an other efficacy endpoint:

2.2.1.3 Other efficacy endpoints

The other efficacy endpoints are defined as:

- Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀
- HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀
- Absolute change from Baseline in International Hidradenitis Suppurativa Severity Score System (IHS4)
- Change from Baseline in the HS-Physician's Global Assessment 6-point scale
- Absolute and percentage change from Baseline in high-sensitivity C-reactive protein (hs-CRP)
- Initiation of systemic antibiotic rescue therapy
- HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ at both Weeks 16 and 48
- Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders
- Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₅₀ during the Maintenance Treatment Period
- Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period
- Change and percentage change from Baseline in lesion counts (abscess count, inflammatory nodule count, AN count, and draining tunnel count)
- AN count of 0, 1, or 2

- AN₂₅, AN₅₀, AN₇₅, AN₉₀, AN₁₀₀ (defined as a 25%, 50%, 75%, 90%, 100% reduction in the total AN count relative to Baseline)
- **Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16**
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48
- Time to flare from Weeks 0 to 16
- Time to flare from Week 16 to 48
- Absolute and percentage change (worst and average skin pain) from Baseline in HS Skin Pain score (11-point numeric rating scale)
- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline
- Absolute change from Baseline in DLQI Total Score
- DLQI Total Score of 0 or 1
- Minimum clinically important difference (MCID) (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4
- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) and Total score
- Patient Global Impression of HS Severity (PGI-S-HS)
- Patient Global Impression of Change of HS Severity (PGI-C-HS)
- Patient Global Impression of Severity of Skin Pain (PGI-S-SP)
- Patient Global Impression of Change of Skin Pain (PGI-C-SP)
- Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor
- Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor

- Responses to the European Quality of Life-5 Dimensions-3 Level questionnaire (EQ-5D-3L), absolute and changes from Baseline in EQ-5D-3L visual analog scale (VAS) scores
- Absolute change from Baseline in Work Productivity and Activity Impairment Questionnaire-Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores
- Domain Scores (effectiveness, convenience and global satisfaction) on the Treatment Satisfaction Questionnaire – Medication-9 (TSQM-9)

Change #4

Flare was removed from the table of power calculations:

Table 2–1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6
Flare ^b	0.99	0.99	Proportion of participants with flare by Week 16=0.09	Proportion of participants with flare by Week 16=0.19	Proportion of participants with flare by Week 16=0.52
Worst Skin Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Skin Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0003 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms

^b Within-participant average of Worst Skin Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Skin Pain score at or above 3 (ie, the threshold for clinically meaningful change from Baseline).

Change #5

The language in **Section 4.2.2.3 Tipping Point Analysis** was updated:

For tipping point analyses, data for participants after the intercurrent event date (See Section 3.9) will be changed to missing prior to imputation ~~but~~ **and, for the bimekizumab treated participants, will not** be changed to non-response after imputation.

Change #6

The language in **Section 4.2.3 Rationale for estimand** was updated:

- A composite estimand strategy will be used for the primary analysis of the primary and binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, ~~flare~~, HS worst skin pain response).

Change #7

Flare was removed from the testing procedure to control for multiplicity:

4.5 Multiple comparisons/multiplicity

To control the overall type I error rate at 0.05 for the multiple comparisons in the primary and secondary efficacy endpoints, a closed testing procedure under a parallel gatekeeping framework will be applied (Sun, 2018).

Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$). Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

1. Step 1: Test HiSCR₅₀ at significance level 0.025.
2. Steps 2 to 56 – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown in Figure 4-1, moving to the next step only if significance achieved at 0.025.
3. In the event that Step 56 is significant at 0.025 for a given dose, then Steps 1 to 56 will be repeated for the other dose using a significance level of 0.05.

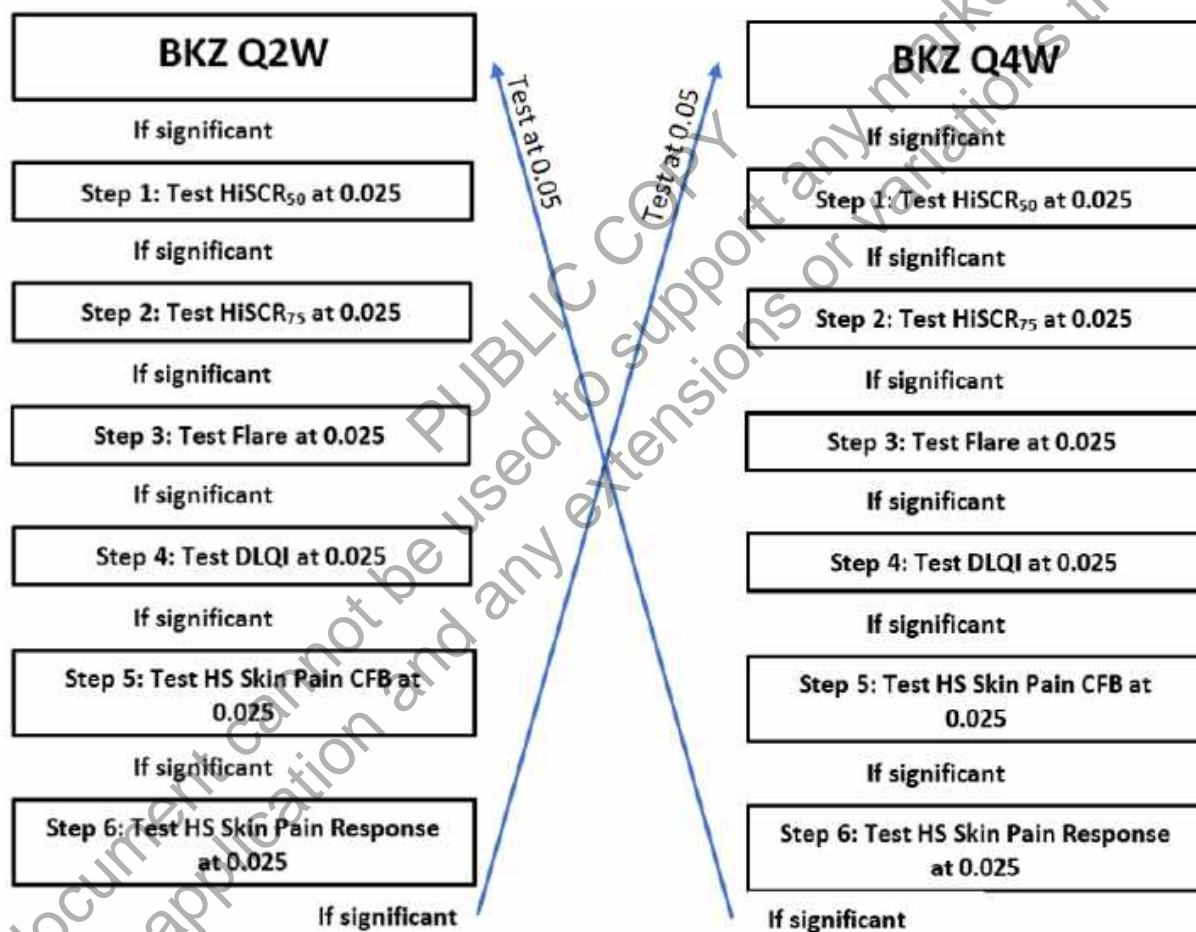
The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment using the analysis methods specified in Section 8.3:

1. Proportion of study participants who achieve HiSCR₇₅ at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
- ~~2. Proportion of study participants who experience at least 1 flare by Week 16, with flare defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline.~~
 - ~~a. bimekizumab 320mg Q2W vs placebo~~
 - ~~b. bimekizumab 320mg Q4W vs placebo~~
2. Absolute CFB in DLQI Total Score at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo

3. Absolute CFB in Skin Pain Score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HSSDD.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
4. Skin pain response at Week 16, based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo

Change #8

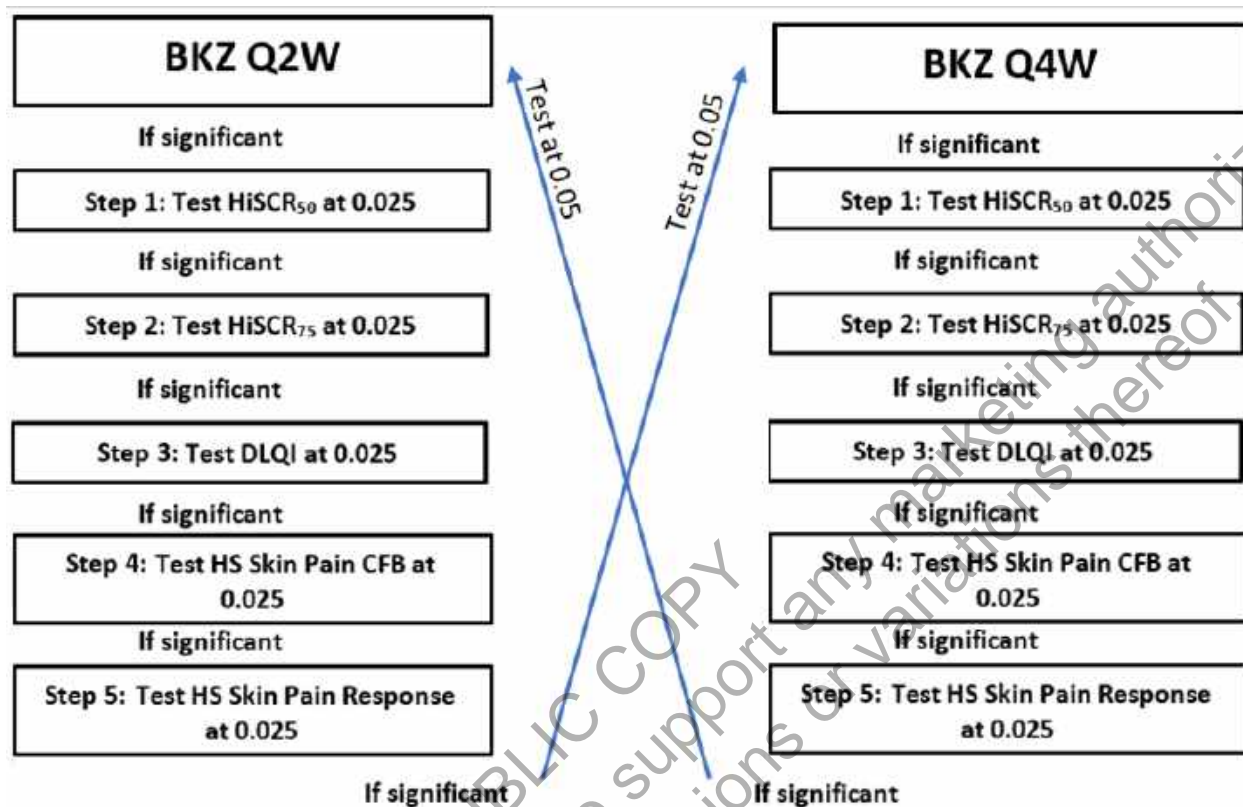
The sequence of testing was changed from the 6-Step schematic below:



AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks
HS skin pain response is tested among study participants with a score of ≥ 3 at Baseline.

To the 5-Step schematic below, which reflects the change to the statistical hierarchy:

Figure 4-1: Sequence of testing



AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks
HS skin pain response is tested among study participants with a score of ≥ 3 at Baseline.

Change #9

The language in **Section 4.8 Examination of subgroups** was updated as follows:

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, ~~flare~~, and worst skin pain response endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period. Additional subgroup analyses will be performed on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, ~~flare~~, and skin pain response endpoints which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

Change #10

Flare was removed from **Table 8-2: Estimand Details and Attributes for Secondary Analyses** as follows:

Table 8-2: Estimand Details and Attributes for Secondary Analyses

		Estimands for Secondary Endpoints			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Secondary Objective: Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS					
HiSCR ₇₅	Secondary (Section 8.3.1)	HiSCR ₇₅ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.
Flare	Secondary (Section 8.3.2)	Flare by Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (flare).	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.

Table 8-2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
DLQI	Secondary (Section 8.3.2.1)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.
DLQI	Secondary - Sensitivity (Section 8.3.2.2)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a DLQI total score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.

Table 8-2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.3.1)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.
HSSDD	Secondary (Section 8.3.3.2)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.

Table 8-2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.4.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.
HSSDD	Secondary Sensitivity (Section 8.3.4.2.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.

Table 8-2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary Sensitivity (Section 8.3.4.2.2)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

AE=adverse event; ANCOVA=analysis of covariance; DLQI=Dermatology Life Quality Index; HiSCR=Hidradenitis Suppurativa Clinical Response; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; IES=intercurrent event(s) strategy; LSMD=Least Squares Mean Difference; MCMC=Markov Chain Monte Carlo; MI=multiple imputation; PLS=Population-level summary; Pop=Population; RS=Randomized Set

^a Analysis includes all study participants in the RS with a Baseline HSSDD Worst Skin Pain score of 3 or higher.

Change #11

Section 8.3.2 Flare by Week 16 was removed, and re-inserted as **Section 8.4.3 Flare by Week 16**. The number of surrounding sections was subsequently updated.

Change #12

Section 8.5 Flare by Week 16 was added to the list of nominal tests.

- HiSCR₅₀ at Week 12
- HiSCR₇₅ at Week 12
- HiSCR₉₀
- HiSCR₁₀₀
- Flare by Week 12
- Flare by Week 16

-
- Time to flare by Week 12 (based on time-to-event analysis per Section 8.4.5 and adjusted appropriately)
 - IHS4 change from Baseline
 - IHS4 percentage change from Baseline
 - HS Physician's Global Assessment: rate of participants who are Clear or Mild
 - DLQI total score change from Baseline at Week 12
 - Worst Skin Pain per HSSDD change from Baseline at Week 12
 - Skin Pain response per HSSDD at Week 12

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

This document has been reviewed and approved per the Review and Approval of Clinical Documents Standard Operating Procedures. Signatures indicate that the final version of the Statistical Analysis Plan (SAP) or amended SAP is released for execution.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Approval Signatures

Name: hs0003-sap-amend-3

Version: 1. 0

Document Number: CLIN-000204829

Title: hs0003-sap-amend-3

Approved Date: 07 Oct 2022

Document Approvals	
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 07-Oct-2022 14:23:43 GMT+0000
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 07-Oct-2022 16:26:25 GMT+0000

STATISTICAL ANALYSIS PLAN

Study: HS0004

Product: Bimekizumab

A PHASE 3, RANDOMIZED, DOUBLE-BLIND, PLACEBO-CONTROLLED,
MULTICENTER STUDY EVALUATING THE
EFFICACY AND SAFETY OF BIMEKIZUMAB IN STUDY
PARTICIPANTS WITH MODERATE TO SEVERE HIDRADENITIS SUPPURATIVA

SAP/Amendment Number	Date
Final SAP	29 May 2020
SAP Amendment 1	15 Oct 2021
SAP Amendment 2	06 Jul 2022

Confidentiality Statement

Confidential

This document is the property of UCB and may not – in full or in part – be passed on, reproduced, published, or otherwise used without the express permission of UCB.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS	9
1 INTRODUCTION	13
2 PROTOCOL SUMMARY	13
2.1 Study objectives	13
2.1.1 Primary objective	13
2.1.2 Secondary objectives	13
2.1.3 Other objectives	13
2.1.4 Exploratory objective	13
2.2 Study endpoints	14
2.2.1 Efficacy endpoints	14
2.2.1.1 Primary efficacy endpoint	14
2.2.1.2 Secondary efficacy endpoints	14
2.2.1.3 Other efficacy endpoints	14
2.2.2 Pharmacokinetic and pharmacogenomic endpoints	16
2.2.2.1 Pharmacokinetic endpoints	16
2.2.2.2 Exploratory pharmacogenomic endpoints	16
2.2.3 Safety endpoints	16
2.2.3.1 Secondary safety endpoints	16
2.2.3.2 Other safety endpoints	16
2.2.4 Immunological endpoints	17
2.3 Study design and conduct	17
2.3.1 Study description	17
2.3.2 Study periods	18
2.3.2.1 Screening Period	18
2.3.2.2 Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)	18
2.3.2.3 Safety Follow-up Visit	18
2.3.3 Study duration per participant	18
2.3.4 Planned number of participants and sites	19
2.3.5 Anticipated regions and countries	19
2.4 Determination of sample size	19
3 DATA ANALYSIS CONSIDERATIONS	21
3.1 General presentation of summaries and analyses	21
3.2 General study level definitions	22
3.2.1 Relative day	22
3.2.2 Mapping of data from Premature End of Treatment visits	23
3.3 Definition of Baseline values	23

3.4	Protocol deviations.....	24
3.5	Analysis sets.....	24
3.5.1	Enrolled Set	24
3.5.2	Randomized Set	24
3.5.3	Safety Set	24
3.5.4	Full Analysis Set.....	24
3.5.5	Active Medication Set	25
3.5.6	Maintenance Set.....	25
3.5.7	Per-Protocol Set	25
3.5.8	Pharmacokinetics Per-Protocol Set.....	25
3.5.9	COVID-19 Free Set	25
3.6	Treatment assignment and treatment groups	25
3.7	Center pooling strategy	25
3.8	Coding dictionaries	26
3.9	Definition of an intercurrent event.....	26
3.10	Changes to protocol-defined analyses	27
4	STATISTICAL/ANALYTICAL ISSUES	28
4.1	Adjustments for covariates	28
4.2	Handling of dropouts or missing data.....	28
4.2.1	Efficacy data	28
4.2.1.1	Handling missing data for the primary efficacy endpoint	28
4.2.1.2	Handling missing data for the secondary efficacy endpoints.....	29
4.2.1.3	Handling missing data for the other efficacy endpoints.....	29
4.2.1.4	Missing Data Overview and Summary.....	29
4.2.2	Missing data algorithms for efficacy analyses.....	31
4.2.2.1	MI – MCMC / Monotone Regression	31
4.2.2.2	MI – MCMC / Reference-based imputation.....	36
4.2.2.3	Tipping Point Analysis	37
4.2.3	Rationale for estimand.....	39
4.2.3.1	Composite estimand	39
4.2.3.2	Hypothetical estimand.....	40
4.2.4	Dates and times	40
4.3	Interim analyses and data monitoring.....	42
4.3.1	Data monitoring committee	42
4.3.2	Interim analysis.....	42
4.4	Multicenter studies.....	42
4.5	Multiple comparisons/multiplicity.....	42
4.6	Use of an efficacy subset of participants	44

4.7	Active-control studies intended to show equivalence.....	44
4.8	Examination of subgroups	44
5	STUDY POPULATION CHARACTERISTICS.....	46
5.1	Study participant disposition.....	46
5.2	Impact of COVID-19	46
5.3	Protocol deviations.....	47
6	DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS	47
6.1	Demographics	47
6.2	Other Baseline characteristics.....	48
6.3	Medical history and concomitant diseases.....	49
6.4	Prior and concomitant medications.....	49
6.4.1	Assignment of medications to study period.....	50
6.4.2	Classification of participants as analgesic, antihistamine users	50
7	MEASUREMENTS OF TREATMENT COMPLIANCE.....	51
8	EFFICACY ANALYSES	51
8.1	Lesion count assessment	52
8.2	Primary efficacy endpoint.....	53
8.2.1	Derivation of HiSCR ₅₀ at Week 16	57
8.2.2	Primary analysis of the primary efficacy endpoint.....	58
8.2.3	Sensitivity analyses of the primary efficacy endpoint.....	59
8.2.3.1	Nonresponse imputation.....	59
8.2.3.2	MI-MCMC / Reference-based imputation	59
8.2.3.3	Tipping point analysis	59
8.2.3.4	Treatment policy.....	60
8.2.3.5	Analysis on observed cases	60
8.2.3.6	Analysis on FAS.....	60
8.2.3.7	Analysis on PPS	60
8.2.3.8	Analysis on CFS	60
8.2.3.9	Analysis including COVID-19 impact as intercurrent event	60
8.2.3.10	Cochran-Mantel-Haenszel test	61
8.2.3.11	Center-by-Treatment Interaction	61
8.3	Secondary efficacy endpoints	62
8.3.1	HiSCR ₇₅ at Week 16.....	67
8.3.2	Flare by Week 16.....	67
8.3.3	Change from Baseline in DLQI Total Score at Week 16	67
8.3.3.1	Primary analysis of change from Baseline in DLQI Total score at Week 16	69
8.3.3.2	Sensitivity analysis of change from Baseline in DLQI Total score at Week 16	69

8.3.4	Change from Baseline in Skin Pain score at Week 16, as assessed by the “worst skin pain” item in the HSSDD	69
8.3.4.1	Primary analysis of change from Baseline in skin pain score at Week 16	70
8.3.4.2	Sensitivity analysis of change from Baseline in skin pain score at Week 16	70
8.3.5	HSSDD skin pain response at Week 16	70
8.3.5.1	Primary analysis of skin pain response at Week 16	70
8.3.5.2	Sensitivity analyses of Skin Pain Response at Week 16	71
8.4	Other efficacy endpoints	71
8.4.1	HiSCR endpoints	71
8.4.1.1	HiSCR ₂₅ , HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀	71
8.4.1.2	Time to response of HiSCR ₂₅ , HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀	72
8.4.1.3	HiSCR ₂₅ , HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀ response at both Weeks 16 and 48	73
8.4.1.4	HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀ among Week 16 Responders	73
8.4.1.5	Time to loss of response of HiSCR ₅₀ , HiSCR ₇₅ , HiSCR ₉₀ , and HiSCR ₁₀₀ in Week 16 responders	73
8.4.1.6	Partial response	74
8.4.2	Lesion count	74
8.4.2.1	Change from Baseline in lesion count	74
8.4.2.2	AN count of 0, 1, or 2	75
8.4.2.3	AN ₂₅ , AN ₅₀ , AN ₇₅ , AN ₉₀ , AN ₁₀₀	75
8.4.3	Flare relative to Baseline	75
8.4.4	Time to flare by Week 16	76
8.4.5	Time to flare by Week 48	76
8.4.6	International Hidradenitis Suppurativa Severity Score System (IHS4)	77
8.4.7	HS-Physician’s Global Assessment 6-point scale	77
8.4.8	High Sensitivity C-Reactive Protein (hs-CRP)	80
8.4.9	Initiation of systemic antibiotic rescue therapy	81
8.4.10	Time to initiation of systemic rescue therapy in the Initial Treatment Period	81
8.4.11	Time to an intercurrent event in the Initial Treatment Period	81
8.4.12	Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)	82
8.4.13	Hidradenitis Suppurativa Symptom Questionnaire (HSSQ)	83
8.4.14	DLQI	84
8.4.15	Hidradenitis Suppurativa Quality of Life (HiSQOL)	84
8.4.16	Patient Global Impression of HS Severity (PGI-S-HS)	86
8.4.17	Patient Global Impression of Change in HS Severity (PGI-C-HS)	86
8.4.18	Patient Global Impression of Severity of Skin Pain (PGI-S-SP)	86

8.4.19	Patient Global Impression of Change in Severity of Skin Pain (PGI-C-SP)	86
8.4.20	Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)	87
8.4.21	Work Productivity and Activity Impairment Questionnaire–Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores.....	87
8.4.22	Treatment Satisfaction Questionnaire – Medication-9	88
8.4.23	Lesion intervention	89
8.5	Additional statistical analyses of other efficacy endpoints	89
9	PHARMACOKINETICS AND PHARMACODYNAMICS	90
9.1	Pharmacokinetics	90
9.2	Pharmacodynamics	91
9.3	Immunogenicity	91
9.3.1	Autoantibodies	91
9.3.2	Anti-bimekizumab antibodies.....	91
10	SAFETY ANALYSES.....	95
10.1	Extent of exposure	95
10.1.1	Exposure during the Initial Treatment Period	96
10.1.1.1	Study medication duration (days).....	96
10.1.1.2	Time at risk (days).....	96
10.1.2	Exposure during the Maintenance Treatment Period	97
10.1.2.1	Study medication duration (days).....	97
10.1.2.2	Time at risk (days).....	97
10.1.3	Exposure during the Initial and Maintenance Treatment Period	98
10.1.3.1	Study medication duration (days).....	98
10.1.3.2	Time at risk (days).....	98
10.2	Adverse events	99
10.2.1	Data considerations.....	99
10.2.1.1	COVID-19 related considerations	101
10.2.2	AE summaries.....	101
10.2.3	Other Safety topics of interest	103
10.2.3.1	Infections (serious, opportunistic, fungal and TB).....	103
10.2.3.2	Malignancies	106
10.2.3.3	Major adverse cardiac event.....	106
10.2.3.4	Neutropenia	108
10.2.3.5	Suicidal Ideation and Behavior	108
10.2.3.6	Inflammatory bowel disease.....	109
10.2.3.7	Hypersensitivity (including anaphylaxis).....	110
10.2.3.8	Hepatic events and PDILI	111
10.3	Clinical laboratory evaluations	111

10.4	Vital signs, physical findings, and other observations related to safety	114
10.4.1	Vital signs	114
10.4.2	Electrocardiograms	115
10.4.3	Other safety endpoints	115
10.4.3.1	Physical examination.....	116
10.4.3.2	Columbia-Suicide Severity Rating Scale (C-SSRS)	116
10.4.3.3	Assessment and management of TB and TB risk factors.....	116
10.4.3.4	Pregnancy testing	117
10.4.3.5	Patient Health Questionnaire (PHQ)-9 scores.....	117
11	REFERENCES	118
12	APPENDICES	120
12.1	Appendix A: MedDRA algorithmic approach to anaphylaxis.....	120
12.2	Appendix B: Definition of CTCAE grades.....	122
13	AMENDMENT TO THE STATISTICAL ANALYSIS PLAN	124
13.1	Amendment 1.....	124
13.2	Amendment 2.....	177
	STATISTICAL ANALYSIS PLAN SIGNATURE PAGE.....	206

LIST OF TABLES

Table 2-1:	Power calculation assumptions and methods.....	20
Table 3-1:	Geographic regions and corresponding countries.....	25
Table 4-1:	Missing data handling approach by endpoint priority and type.....	30
Table 4-2:	Imputation allowable ranges by variable	33
Table 8-1:	Estimand Details and Attributes for Primary Endpoint	53
Table 8-2:	Estimand Details and Attributes for Secondary Analyses	62
Table 8-3:	Dermatology Life Quality Index.....	68
Table 8-4:	HS-Physician's Global Assessment 6-point scale	77
Table 8-5:	Hidradenitis Suppurative Quality of Life	85
Table 10-1:	Cardiovascular event classifications.....	107
Table 10-2:	Neuropsychiatric event classifications.....	109
Table 10-3:	IBD event classifications	110
Table 10-4:	Definitions of Markedly Abnormal Liver Function Values	112
Table 10-5:	Definitions of Markedly Abnormal Biochemistry Values.....	113
Table 10-6:	Definitions of Markedly Abnormal Hematology Values.....	113
Table 10-7:	Definitions of Markedly Abnormal Blood Pressure Values	115

Table 12–1:	Definition of CTCAE grades by biochemistry parameters	122
Table 12–2:	Definitions of CTCAE grades by hematology parameter	122
Table 2–1:	Power calculation assumptions and methods.....	180
Table 13–2:	Imputation allowable ranges by variable	182
Table 8–1:	Estimand Details and Attributes for Primary Endpoint	187
Table 813–2:	Estimand Details and Attributes for Secondary Analyses	189
Table 13–1:	Cardiovascular event classifications	201
Table 13–2:	Neuropsychiatric event classifications.....	203
Table 13–3:	IBD event classifications	204

LIST OF FIGURES

Figure 2-1:	Study Schematic.....	17
Figure 4-1:	Sequence of testing	44
Figure 13-1:	Sequence of testing	185

LIST OF ABBREVIATIONS

List of Abbreviations

%ΔAN	percentage change from Baseline in abscess and inflammatory nodule count
ADAb	anti-bimekizumab antibodies
AE(s)	adverse event(s)
ALP	alkaline phosphatase
ALQ	above the limit of quantification
ALT	alanine aminotransferase
AMS	Active Medication Set
AN	abscess and inflammatory nodule
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
BKZ	bimekizumab
BLQ	below the limit of quantification
CFB	change from Baseline
CFS	COVID-19 Free Set
CI	confidence interval
CMH	Cochran-Mantel-Haenszel
COVID-19	coronavirus disease 2019
CSR	clinical study report
CTCAE	Common Terminology Criteria for Adverse Events
CV	coefficient of variation
CV-CAC	Cardiovascular Event Adjudication Committee
DLQI	Dermatology Life Quality Index
DMC	data monitoring committee
eCDF	empirical cumulative distribution function
ECG	electrocardiogram
eCRF	electronic case report form
eC-SSRS	electronic Columbia-Suicide Severity Rating Scale
EAER	exposure adjusted event rate
EAIR	exposure adjusted incident rate

List of Abbreviations

EQ-5D-3L	European Quality of Life-5 Dimensions-3 Level questionnaire
ES	Enrolled Set
FAS	Full Analysis Set
FDA	Food and Drug Administration
geoCV	geometric coefficient of variation
GGT	gamma-glutamyltransferase
HiSCR	Hidradenitis Suppurativa Clinical Response
HiSQOL	Hidradenitis Suppurativa Quality of Life
HLT	high level term
HS	hidradenitis suppurativa
hs-CRP	high sensitivity C-reactive protein
HSSDD	Hidradenitis Suppurativa Symptom Daily Diary
HSSQ	Hidradenitis Suppurativa Symptom Questionnaire
IBD	Inflammatory bowel disease
IBD-CAC	Inflammatory Bowel Disease Adjudication Committee
ICF	Informed Consent Form
ICH	International Council for Harmonisation
IDC	Infectious Disease Committee
IGRA	interferon gamma release assay
IHS4	International Hidradenitis Suppurativa Severity Scoring System
IMP	investigational medicinal product
LFT	liver function tests
LLOQ	lower limit of quantification
LOCF	last observation carried forward
LSM	least square mean
MACE	major cardiovascular events
MAR	missing at random
MCID	minimal clinically important difference
MCMC	Markov-Chain Monte Carlo
MedDRA	Medical Dictionary for Regulatory Activities
MI	multiple imputation
MI-MCMC	multiple imputation Markov-Chain Monte Carlo

List of Abbreviations

MS	Maintenance Set
MSR	minimum significant ratio
n	number of study participants
NAb	neutralizing antibody
NI	Negative Immunodepletion
nR	New Ratio
NRI	nonresponder imputation
NRS	numeric rating scale
NS	Negative Screen
OC	observed case
PD	pharmacodynamic(s)
pDILI	potential drug induced liver injury
PEOT	premature end of treatment
PGI-C-HS	Patient Global Impression of Change in Hidradenitis Suppurativa Severity
PGI-C-SP	Patient Global Impression of Change in Severity of Skin Pain
PGI-S-HS	Patient Global Impression of Hidradenitis Suppurativa Severity
PGI-S-SP	Patient Global Impression of Severity of Skin Pain
PHQ-9	Patient Health Questionnaire 9
PI	Positive Immunodepletion
PK	pharmacokinetic(s)
PK-PPS	Pharmacokinetic Per-Protocol Set
PPS	Per-Protocol Set
PS	Positive Screen
PT	preferred term
Q2W	every 2 weeks
Q4W	every 4 weeks
QTcF	QT corrected for heart rate using Fridericia's formula
RS	Randomized Set
SAP	statistical analysis plan
SD	standard deviation
SE	standard error

List of Abbreviations

SFU	Safety Follow-up
SIB	suicidal ideation and behavior
SMQ	standardized MedDRA query
SOC	system organ class
SS	Safety Set
SSD	Safety Signal Detection
TEAE	treatment-emergent adverse event
TEMA	treatment-emergent markedly abnormal
TSQM-9	Treatment Satisfaction Questionnaire – Medication 9
ULN	upper limit of normal
VAS	visual analogue scale
WHODD	World Health Organization Drug Dictionary
WPAI-SHP	Work Productivity and Activity Impairment Questionnaire-Specific Health Problem

1 INTRODUCTION

This statistical analysis plan (SAP) defines the scope of statistical analyses and provides a detailed description of statistical methodology to support the final clinical study report (CSR).

The SAP is based on the Protocol Amendment 4, 6 May 2022 and the Japan-specific amendment 4.1, 16 June 2022. If a future protocol amendment necessitates a substantial change to the statistical analysis of the study data, this SAP may be amended accordingly. The content of this SAP is compatible with the International Council for Harmonisation (ICH)/Food and Drug Administration (FDA) E9 Guidance documents.

2 PROTOCOL SUMMARY

2.1 Study objectives

2.1.1 Primary objective

The primary objective of this study is to evaluate the efficacy of bimekizumab in study participants with moderate to severe hidradenitis suppurativa (HS).

2.1.2 Secondary objectives

The secondary objectives of this study are to:

- Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS
- Evaluate the safety of bimekizumab in study participants with moderate to severe HS

2.1.3 Other objectives

The other objectives of this study are to:

- Evaluate the efficacy of bimekizumab on Hidradenitis Suppurativa Clinical Response (HiSCR), other HS Scores, and other clinical measures of disease activity at various timepoints in study participants with moderate to severe HS
- Evaluate the efficacy of bimekizumab on abscesses, nodules, and draining tunnels at various timepoints in study participants with moderate to severe HS
- Evaluate the efficacy of bimekizumab on patient-reported outcome measures at various timepoints in study participants with moderate to severe HS
- Evaluate the effect of bimekizumab on other safety measures at various timepoints in study participants with moderate to severe HS
- Evaluate the pharmacokinetics (PK) of bimekizumab in study participants with moderate to severe HS
- Evaluate the immunogenicity of bimekizumab (antidrug antibodies) in study participants with moderate to severe HS

2.1.4 Exploratory objective

The exploratory objective of the study is to evaluate biomarkers in study participants with moderate to severe HS.

2.2 Study endpoints

2.2.1 Efficacy endpoints

2.2.1.1 Primary efficacy endpoint

The primary efficacy endpoint is the HiSCR₅₀ (defined as at least a 50% reduction from Baseline in the total abscess and inflammatory nodule [AN] count with no increase from Baseline in abscess or draining tunnel count) at Week 16.

2.2.1.2 Secondary efficacy endpoints

The secondary efficacy endpoints are defined as:

- HiSCR₇₅ response (defined as at least a 75% reduction from Baseline in the total AN count with no increase from Baseline in abscess or draining tunnel count) at Week 16
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16
- Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16
- Absolute change from Baseline (CFB) in Skin Pain score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HS Symptom Daily Diary (HSSDD)
- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline

2.2.1.3 Other efficacy endpoints

The other efficacy endpoints are defined as:

- Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀
- HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀
- Absolute change from Baseline in International Hidradenitis Suppurativa Severity Score System (IHS4)
- Change from Baseline in the HS-Physician’s Global Assessment 6-point scale
- Absolute and percentage change from Baseline in high-sensitivity C-reactive protein (hs-CRP)
- Initiation of systemic antibiotic rescue therapy
- HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ at both Weeks 16 and 48
- Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders
- Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₅₀ during the Maintenance Treatment Period

- Partial responders (defined as a $\geq 25\%$ reduction in AN count from Baseline [Week 0]) at Week 16 who progress to HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period
- Change and percentage change from Baseline in lesion counts (abscess count, inflammatory nodule count, AN count, and draining tunnel count)
- AN count of 0, 1, or 2
- AN₂₅, AN₅₀, AN₇₅, AN₉₀, AN₁₀₀ (defined as a 25%, 50%, 75%, 90%, 100% reduction in the total AN count relative to Baseline)
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48
- Time to flare from Weeks 0 to 16
- Time to flare from Week 16 to 48
- Absolute and percentage change (worst and average skin pain) from Baseline in HS Skin Pain score (11-point numeric rating scale)
- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline
- Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline
- Absolute change from Baseline in DLQI Total Score
- DLQI Total Score of 0 or 1
- Minimum clinically important difference (MCID) (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4
- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) and Total score
- Patient Global Impression of HS Severity (PGI-S-HS)
- Patient Global Impression of Change of HS Severity (PGI-C-HS)
- Patient Global Impression of Severity of Skin Pain (PGI-S-SP)
- Patient Global Impression of Change of Skin Pain (PGI-C-SP)

- Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor
- Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor
- Responses to the European Quality of Life-5 Dimensions-3 Level questionnaire (EQ-5D-3L), absolute and changes from Baseline in EQ-5D-3L visual analog scale (VAS) scores
- Absolute change from Baseline in Work Productivity and Activity Impairment Questionnaire-Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores
- Domain Scores (effectiveness, convenience and global satisfaction) on the Treatment Satisfaction Questionnaire – Medication-9 (TSQM-9)

2.2.2 Pharmacokinetic and pharmacogenomic endpoints

2.2.2.1 Pharmacokinetic endpoints

The PK endpoint is the plasma bimekizumab concentrations over the study duration.

2.2.2.2 Exploratory pharmacogenomic endpoints

[REDACTED]

A specific SAP will be written to describe the analysis methods for those endpoints, as the results will not be summarized in the CSR. The nature and format of these analyses will be detailed in this SAP.

2.2.3 Safety endpoints

2.2.3.1 Secondary safety endpoints

The secondary safety endpoints are

- Treatment-emergent Adverse Events (TEAEs)
- Serious TEAEs
- TEAEs leading to withdrawal from study

2.2.3.2 Other safety endpoints

The other safety endpoints are

- Adverse events of special interest (Hy's Law)
- Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (including anaphylaxis), suicidal ideation and behavior, major adverse cardiovascular events, hepatic events and potential drug-induced liver injury (PDILI), malignancies, and inflammatory bowel disease.

- Absolute change from Baseline in the Patient Health Questionnaire (PHQ-9) score
- Absolute change from Baseline in vital signs
- Absolute change from Baseline in clinical laboratory values (chemistry and hematology)
- Electrocardiogram (ECG) results

2.2.4 Immunological endpoints

The immunological endpoints are

- Bimekizumab antidrug antibodies
- Bimekizumab neutralizing antibodies

The results of the bimekizumab neutralizing antibody analysis will not be summarized in the CSR for this study. All neutralizing antibody analyses will be detailed in the integrated immunogenicity SAP.

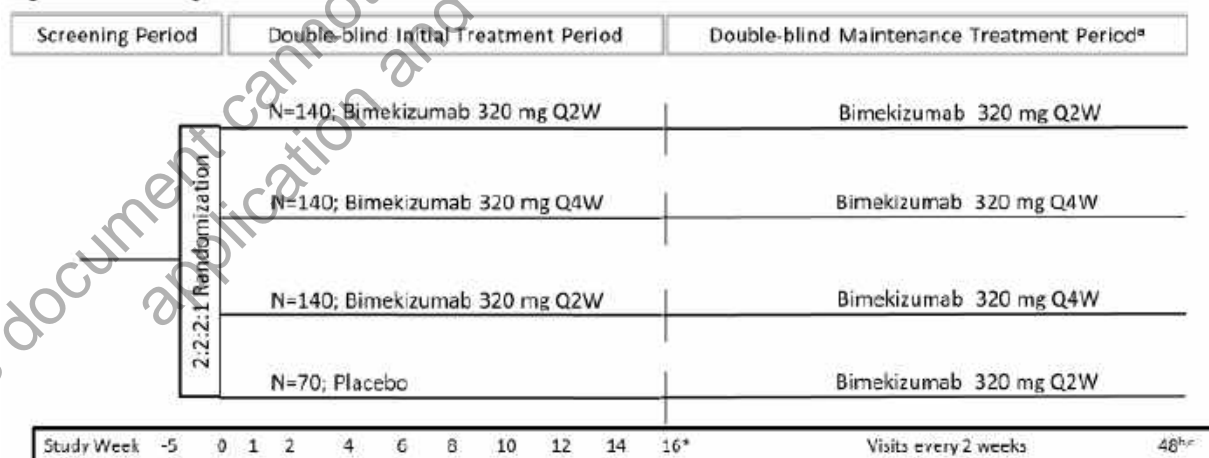
2.3 Study design and conduct

2.3.1 Study description

HS0004 is a Phase 3, randomized, double-blind, placebo-controlled, multicenter, pivotal study evaluating the efficacy and safety of bimekizumab in study participants with moderate to severe HS. Study participants meeting the inclusion criteria who do not meet any exclusion criteria will complete a Screening Period of 14 days to up to 5 weeks; a double-blind, 48-week Treatment Period comprising a 16-week Initial Treatment Period and 32-week Maintenance Treatment Period; and a 20-week Safety Follow-up (SFU) Period following the final injection of investigational medicinal product (IMP) if study participants do not enter a subsequent extension study (HS0005) or withdraw prematurely from treatment.

Study participants will be randomized in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to 1 of 3 dose regimens of bimekizumab or placebo as shown in the schematic (Figure 2-1).

Figure 2-1: Study Schematic



HiSCR₅₀=a 50% reduction in the total abscess and inflammatory nodule count with no increase from Baseline in abscess or draining tunnel count;

IMP=investigational medicinal product; Q2W=every 2 weeks; Q4W=every 4 weeks

*Week 16 = primary endpoint (HiSCR₅₀ bimekizumab versus placebo)

a Study participant should discontinue from the study from Week 32 on if no partial response is achieved (partial response is defined as $\geq 25\%$ improvement in abscess and inflammatory nodule count relative to Baseline [Week 0] lesion values,)

b Study participants achieving an improvement of at least 25 % in abscess and inflammatory nodule count continue in HS0005 (Extension Study).

c 20-week Safety Follow-up (from last IMP injection) for any study participant who discontinues from study prior to Week 48, or who does not continue in HS0005.

2.3.2 Study periods

2.3.2.1 Screening Period

The Screening Period will last a minimum of 14 days up to a maximum of 5 weeks prior to randomization.

2.3.2.2 Initial Treatment Period (Weeks 0-16) and Maintenance Treatment Period (Weeks 16-48)

Eligible study participants will be randomized in a 2:2:2:1 ratio as noted in the Study Schema to:

- Bimekizumab 320mg Q2W from Weeks 0 to 48
- Bimekizumab 320mg Q4W from Weeks 0 to 48
- Bimekizumab 320mg Q2W to Week 16, continuing on 320mg Q4W from Weeks 16 to 48
- Placebo to Week 16, continuing on bimekizumab 320mg Q2W from Weeks 16 to 48

2.3.2.3 Safety Follow-up Visit

All study participants not continuing in the extension study, including those withdrawn from IMP as well as those completing all study visits, will have an SFU Visit 20 weeks after their final dose of IMP.

2.3.3 Study duration per participant

The total duration of study participation in HS0004 will be 68 to 71 weeks for those who complete HS0004 and do not participate in the extension study HS0005 and 50 to 53 weeks for those who participate in HS0005 and, thus, do not participate in the 20-week SFU Period. The study is comprised of the following periods:

- Screening Period: 14 days up to a maximum of 5 weeks prior to randomization
- Initial Treatment Period: 16 weeks
- Maintenance Treatment Period: 32 weeks
- Safety Follow-Up Period: 20 weeks after the last dose of IMP

A study participant will be considered to have completed the study if he or she completed the Week 48 visit.

The end of the study is defined as the date of the last scheduled procedure for the last study participant in the study globally, including the SFU, as applicable.

2.3.4 Planned number of participants and sites

A total of approximately 490 study participants will be randomized into the study. The planned number of study sites is approximately 100.

2.3.5 Anticipated regions and countries

The regions planned for study conduct are Western Europe, Central/Eastern Europe, North America and Asia/Australia, with possible extension to other regions and countries.

A subset of tables, listings, and figures will be repeated for the subgroup of participants randomized in Japan. A list of these outputs will be added to the tables, figures, and listings shells.

2.4 Determination of sample size

A total of 490 study participants will be randomly assigned in a 2:2:2:1 ratio (stratified by Hurley Stage and current antibiotic use) to the following treatment arms:

- Bimekizumab 320mg Q2W during Initial Treatment Period (Weeks 0-16) and Maintenance Treatment (Weeks 16-48) Period, N=140
- Bimekizumab 320mg Q2W during Initial Treatment Period (Weeks 0-16), and Bimekizumab 320mg Q4W during Maintenance Treatment Period (Weeks 16-48), N=140
- Bimekizumab 320mg Q4W during Initial Treatment (Weeks 0-16) and Maintenance Treatment Periods (Weeks 16-48), N=140
- Placebo during Initial Treatment Period (Weeks 0-16), and Bimekizumab 320mg Q2W during Maintenance Treatment Period (Weeks 16-48), N=70

The analysis of the primary efficacy endpoint and secondary efficacy endpoints are based on a comparison of bimekizumab versus placebo at Week 16, with alpha adjustment strategy as indicated in Section 4.5.

The power to detect a statistically significant difference for each of the endpoints are shown in Table 2-1. Notably, with a 2-sided significance level of 0.025, the sample size of 140:70 provides 73% power for detecting at least a difference of 1.5 (bimekizumab Q4W vs placebo) for the Worst Skin Pain change from Baseline (CFB) endpoint.

Given the high level of power for each of the primary and key secondary endpoints at the 0.025 significance level for the Q2W comparison (power \geq 0.89), and per the alpha spending strategy, there is a high likelihood that the Q4W comparison of Worst Skin Pain CFB vs placebo will be allowed to be tested against the 0.05 level of significance. The power for this latter test is 81%. The sample size is thus ultimately driven by the Worst Skin Pain CFB endpoint.

Furthermore, the randomization ratio of 2:2:2:1 has been chosen to provide study participants with a high probability (6/7 ~ 86%) of being randomized to active study drug.

After randomization for this study was complete, an additional endpoint to assess Worst Skin Pain response was included in the sequential testing procedure. This additional endpoint is based on the threshold for clinically meaningful change and is defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score at Week 16 among study participants with a score of ≥ 3 at Baseline. Note that the power calculations reported in Table 2-1 for this

endpoint are based on the sample size that was initially driven by the Worst Pain CFB endpoint as described above. With a 2-sided significance level of 0.025, the sample size of 104:52 in the subset of participants reporting Baseline HSSDD worst skin pain score at or beyond the threshold for clinically meaningful change (ie, Baseline HSSDD ≥ 3) provides 53% power for detecting a statistically significant difference between bimekizumab Q4W and placebo in the proportion of Worst Pain responders.

Note that the power at the 0.025 level of significance associated with this endpoint for the comparison between bimekizumab Q2W and placebo is 95%. The Q4W comparison of Worst Skin Pain response vs placebo against the 0.05 level of significance is therefore likely, and the power at this significance level is 65%. Given the strength of the power for the bimekizumab Q2W arm vs placebo across endpoints, it is considered acceptable to have a relatively low power for this final endpoint in the testing sequence for the bimekizumab Q4W treatment arm vs placebo.

Table 2–1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
Flare	0.99	0.99	Proportion of participants with flare by Week 16=0.09	Proportion of participants with flare by Week 16=0.19	Proportion of participants with flare by Week 16=0.52
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6
Worst Skin Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Skin Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0004 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms

^b Within-participant average of Worst Skin Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Skin Pain score at or above 3 (ie, the threshold for clinically meaningful change from Baseline)

3 DATA ANALYSIS CONSIDERATIONS

3.1 General presentation of summaries and analyses

Statistical analysis and generation of tables, figures, study participants data listings, and statistical output will be performed using SAS Version 9.4 or higher. All tables and listings will use Courier New font size 9.

Descriptive statistics will be displayed to provide an overview of the study results. For continuous variables, descriptive statistics will include number of study participants with available measurements (n), mean, standard deviation (SD), median, minimum, and maximum.

For PRO continuous variables, descriptive statistics will also include variable score, absolute and percentage changes from baseline, Q1 and Q3, 10th, and 90th percentiles.

If no participants have data at a given time point, then only n=0 will be presented. The other descriptive statistics will be left blank. If $n < 3$ then the n, minimum, and maximum only will be presented. The other descriptive statistics will be left blank. If $n = 3$ n, mean, median, minimum and maximum will be presented only. The other descriptive statistics will be left blank.

For categorical variables, the number and percentage of study participants in each category will be presented. Unless otherwise noted, the denominator for percentages will be based on the number of study participants included in the respective analysis set. Study participants with missing data will be accounted for using the following approaches:

- For summaries of demographics and Baseline characteristics: percentages will be based on all study participants in the analysis set and a “Missing” category (corresponding to study participants with missing data for the variable being summarized) will be included as the last row in the list of categories being summarized.
- For summaries of efficacy and safety endpoints, unless otherwise specified: percentages will be based only on those study participants with observed data for the variable being summarized. As the denominator may be different from the number of study participants in the analysis set being considered, the denominator will be displayed in the table. The general format for displaying this will be “n/Nsub (%)”.

Percentages will be presented to 1 decimal place. If the percentage is 100%, a decimal will not be presented. If the count is 0, the percentage will not be presented. Typically, the % sign will be presented in the column header, but not with each individual value.

For the purpose of the tabulations the lower and upper confidence limits for the percentages will be truncated at 0 and 100% respectively. Confidence intervals (CIs) for the response rates in efficacy summaries based on nonresponder imputation (NRI) will be computed using the Wilson approximation.

For bimekizumab PK concentrations, summary statistics will include geometric mean, geometric coefficient of variation (CV), 95% CIs for geometric mean, arithmetic mean, SD, median, minimum, and maximum. All summaries of PK variables will be based on the observed values. No imputation will be used.

Decimal places for descriptive statistics will be subject to the following rules:

- “n” will be an integer

- Mean, SD, and median will use 1 additional decimal place compared to the original data
- CV [%] will be presented with 1 decimal place
- Minimum and maximum will have the same number of decimal places as the original value.

Derived variables in general will display the mean, SD, and median to 1 more decimal place than the variables used in the derivation. If the number of decimal places reported in the raw data is varied then use either the maximum raw number of reported decimal places or 3, whichever is the lowest, as a guide for the descriptive statistics.

When reporting individual values and descriptive statistics for PK concentration data, the following rules will apply with regard to rounding and precision:

- Individual values will be reported to the same level of precision as received from the bioanalytical laboratory
- Descriptive statistics will be reported to the same level of precision as the individual data for the minimum and maximum, and to 1 additional significant figure for the mean (arithmetic and geometric), median, SD, and 95% CI for the geometric mean
- The geometric coefficient of variances (geoCV) will be reported as a percentage to 1 decimal place

Statistical tests of efficacy variables will be presented as 2-sided p-values rounded to 3 decimal places. P-values less than 0.001 will be presented as “<0.001” and p-values greater than 0.999 will be presented as “>0.999.” Statistical comparisons will be 2-sided and will be performed at the 0.05 level of significance unless specified otherwise. The significance levels used as part of the multiple testing procedure are detailed in Section 4.5.

Per protocol, visit windows are ± 3 days from the date of first dose. The 20-week SFU Visit window is ± 7 days from the date of the final dose. All by-visit summaries will contain nominal (ie, scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for assessments that occur within a 3-day time window of a scheduled visit. In that case, the assessment will be mapped to the corresponding scheduled visit and will be used for the analysis. This will only occur for selected vendor data. The only exception to this rule is for unscheduled assessments that occur up to 3 days after the Baseline visit. These unscheduled visits will remain as unscheduled as the Baseline assessment cannot be after the first dose of study drug administration. See Section 3.3 for more details on the definition of Baseline values.

A complete set of data listings containing all documented data as well as calculated data (eg, change from Baseline) will be generated.

3.2 General study level definitions

3.2.1 Relative day

The relative day will be included in different listings and will be calculated as follows:

- If the start (stop) date occurred on or after the first dose, but prior to the double-blind drug stop date, relative day is calculated as start (stop) date minus first dose date + 1

- If the start (stop) date occurred after the last dose of double-blind drug, the relative day to the most recent dose is calculated as start (stop) date minus most recent dose date. The relative day in this situation will be preceded by a '+'
- If the start (stop) date occurred before the first dose, the relative day is calculated as start (stop) date minus first dose date. The relative day in this situation will be preceded by a '-'.

For AEs, relative days for start and stop dates will be calculated as the number of days since the first injection of the medication. For non-treatment emergent AEs, relative day of onset will be negative if the event started and stopped before the first dose. Relative day will only be computed for fully completed dates and will be missing for partial dates.

Relative day will be calculated from first dose of IMP for all treatment groups, and additionally from first dose of bimekizumab for the Placebo/BKZ 320mg Q2W arm.

3.2.2 Mapping of data from Premature End of Treatment visits

If the Premature end of treatment (PEOT) visit occurs at a scheduled visit as outlined in the schedule of assessments, then no mapping is necessary, and any early withdrawal assessments will correspond to that scheduled visit. The PEOT assessments that occur on a date after a scheduled visit will be assigned to the next scheduled site visit per the protocol following the last visit where assessments were available. This approach means that there is a chance that data will be mapped to a visit where a given assessment was not actually collected per the protocol schedule of assessments. Such data will not be summarized in by-visit tables (though it will be available in the listings).

The only exception to the above rule is for anti-bimekizumab antibody assessments, in which all PEOT assessments will be assigned to the next scheduled visit at which anti-bimekizumab antibodies are assessed. All by-visit summaries will contain nominal visits only. Unscheduled visits will not be mapped to scheduled visits. Note that based on the early withdrawal mapping conventions described above, a mapped PEOT visit is considered as observed at that visit and will be summarized as such in the tables.

3.3 Definition of Baseline values

Section 8.3.4 details the derivation of the Baseline value for the HSSDD assessment. For all other assessments, the below applies.

A Baseline value for a participant is defined as the latest non-missing measurement for that participant up to and including the day of administration of first study medication, unless otherwise stated. If a Baseline assessment is taken on the same day as first administration of study medication, it is eligible to be used as the Baseline value, even in the case that the time of the assessment is recorded as taking place after the time of first study medication administration. This is considered acceptable as this measurement is still the best representation of the Baseline value of the given assessment since it is highly unlikely that the study medication could have an impact on any measurement in such a short period of time. However, such cases should be rare as study center personnel are instructed to do all assessments at the Baseline visit prior to administering study medication. One exception to this rule is plasma concentration of bimekizumab. If Baseline plasma concentration is measured at a time after the first administration of study medication, then it will not be eligible to be considered as a Baseline plasma concentration. Such cases will be discussed with the quantitative clinical pharmacologist.

For randomized participants for whom no start date of treatment is available, the Baseline value will be considered as the last available value on or before the randomization date.

If a Baseline measurement is missing or not collected, and a Screening value is available, the Screening value will be utilized as Baseline instead.

Baseline values for component scores should be computed using components from the same visit where the relevant measurements were recorded prior to dosing. For example, if the Screening visit has all of the components, but the Baseline visit is missing 1 or more components, the Baseline value for the component score should be calculated using the Screening visit values.

When the time of first dose is derived, it will be based on the first injection of study treatment, regardless of whether or not it is an active treatment.

3.4 Protocol deviations

Important protocol deviations are deviations from the protocol that could potentially have a meaningful impact on study conduct or on the primary and key secondary efficacy, key safety, or PK outcomes for an individual participant. The criteria for identifying important protocol deviations will be defined within the appropriate protocol-specific document at study start. Important protocol deviations will be reviewed as part of the ongoing data cleaning process. Important protocol deviations including those that lead to exclusion from the analysis sets will be identified and documented prior to unblinding.

Deviations related to the Coronavirus Disease 2019 (COVID-19) global pandemic are unavoidable deviations from the protocol due to confirmed COVID-19 infection, suspected COVID-19 infection, general circumstances around COVID-19 without infection or any other deviation from the protocol due to COVID-19. COVID-19 protocol deviations will also be reviewed separately as part of the ongoing data cleaning process.

3.5 Analysis sets

3.5.1 Enrolled Set

The Enrolled Set (ES) will consist of all participants who have given informed consent.

3.5.2 Randomized Set

The Randomized Set (RS) will consist of all participants randomized into the study.

3.5.3 Safety Set

The Safety Set (SS) will consist of all participants who received at least 1 dose (full or partial) of IMP. The SS will be used for the demographic, safety, and immunogenicity analyses.

3.5.4 Full Analysis Set

The Full Analysis Set (FAS) will consist of all study participants who received at least 1 dose (full or partial) of IMP and had a valid Baseline measurement and a post-Baseline measurement for abscess, inflammatory nodules, and draining tunnel counts.

3.5.5 Active Medication Set

The Active Medication Set (AMS) will consist of all participants who have received at least 1 dose (full or partial) of bimekizumab. The AMS will be used for summaries of safety that include all data from the Initial Treatment Period and/or Maintenance Treatment Period.

3.5.6 Maintenance Set

The Maintenance Set (MS) will consist of all participants who have received at least 1 dose (full or partial) of bimekizumab in the Maintenance Treatment Period.

3.5.7 Per-Protocol Set

The Per-Protocol Set (PPS) will consist of all study participants in the FAS who had no important protocol deviations affecting the primary efficacy variable. Important protocol deviations will be predefined and study participants with important protocol deviations will be evaluated during ongoing data cleaning meetings prior to unblinding of the data.

3.5.8 Pharmacokinetics Per-Protocol Set

The Pharmacokinetics Per-Protocol Set (PK-PPS) will consist of study participants who received at least 1 full dose of bimekizumab and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the PK.

3.5.9 COVID-19 Free Set

The COVID-19 Free Set (CFS) will consist of all study participants randomized into the study and who have no COVID-19 impact through Week 16. This analysis set will be used for sensitivity analysis of the primary efficacy endpoint.

3.6 Treatment assignment and treatment groups

It is expected that participants receive treatment as randomized and hence safety analyses will be based on the SS, as randomized. However, if after unblinding it is determined that participants randomized to placebo in the Initial Treatment Period received bimekizumab at any time within the first 16 weeks, then for safety analyses these participants will be reallocated to the appropriate bimekizumab treatment group, unless otherwise specified. Participants randomized to bimekizumab will only be reallocated to the placebo treatment group if they never received bimekizumab. Efficacy analyses will be according to randomized treatment and not actual treatment received.

For the purposes of Initial Treatment Period analyses for the 320mg Q2W dosing regimen, the bimekizumab treatment arms of 320mg Q2W/Q2W and bimekizumab 320mg Q2W/Q4W treatment groups will be pooled.

3.7 Center pooling strategy

Geographic regions have been categorized as North America, Western Europe, Central/Eastern Europe, and Asia/Australia. Below is a table of geographic regions with corresponding countries.

Table 3–1: Geographic regions and corresponding countries

Region	Countries
North America	Canada, United States

Table 3–1: Geographic regions and corresponding countries

Region	Countries
Western Europe	France, Germany, Ireland, Italy, Spain, United Kingdom
Central/Eastern Europe	Bulgaria, Czech Republic, Hungary, Poland
Asia/Australia	Australia, Israel, Japan

The following center pooling algorithm will be used for each geographic region:

- If a center has 21 or more participants, then no pooling will be done for that center.
- Centers with fewer than 21 participants will be ordered from largest to smallest with pooling proceeding in the following manner:
 - Two or more centers will be combined until the cumulative participant total is at least 21.
 - Once a pooled center has at least 21 participants, the process will continue in an iterative fashion for the subsequent centers in the ordered list, where a new pooled center begins each time at least 21 participants has been reached in the previous pool.
 - If this iterative process reaches the end of the ordered list of centers where the final pooled center has fewer than 21 participants, then the participants from the centers in that pool will be combined with the pooled center formed in the previous iteration.

This procedure is only to be performed within a geographic region – there will be no pooling of centers across regions.

In the event that the percentage of randomized participants is less than 10% in either of the Asia/Australia or Central/Eastern Europe regions, the two regions will be combined as a geographic region stratum for efficacy modeling, so that there are no modeling convergence issues across efficacy variables.

3.8 Coding dictionaries

Adverse events and medical history will be coded using the Medical Dictionary for Regulatory Activities (MedDRA®) version 19.0.

Medications will be coded according to the World Health Organization Drug Dictionary (WHODD) version MAR2021 B3 or later. Medical procedures will not be coded.

3.9 Definition of an intercurrent event

Handling of intercurrent events is one of the key elements for the analysis of efficacy endpoints.

An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy (See Section 8.2.2).

Receipt of systemic antibiotic rescue medication is defined as initiating any systemic antibiotic on or after Baseline for any reason (including in response to an AE). The only exception to this rule is if a participant randomized to the antibiotic stratum on a tetracycline antibiotic interrupts

their stable dose of tetracycline antibiotic during the study and subsequently restarts the same tetracycline antibiotic as confirmed using the coded preferred term. The restarted dose and frequency of the antibiotic must be the same or lower than the regimen prior to the interruption.

The dates of an intercurrent event are as follows:

- For receipt of systemic antibiotic rescue medication: start date of the antibiotic
- For discontinuation of study treatment due to an AE or lack of efficacy: Last study treatment date + 17 days. Note: study treatment discontinuation includes study discontinuation.

The choice of 17 days is intended to capture the interval between dosing and lesion assessments (14 days), as well as the visit window (3 days).

An additional sensitivity analysis will be conducted where missing data due to COVID-19 will be considered an intercurrent event and will be imputed as a nonresponse at that particular visit. This will be identified when there are missing data at a visit that has been impacted by COVID-19 according to the COVID-19 impact CRF page. The date of this intercurrent event will be the date of the impacted visit.

3.10 Changes to protocol-defined analyses

The MS and AMS were added as analysis sets.

The endpoints for PGI-S-HS, PGI-S-SP, PGI-C-HS, and PGI-C-SP were clarified in Section 2.2.1.3 to indicate that absolute change from Baseline will not be calculated, as these are categorical endpoints.

The HiSQOL endpoint was clarified to show that there are only 3 domains: symptoms, psychosocial, activities and adaptations and to add total score.

In Protocol Amendment 4, the secondary efficacy endpoint for skin pain response based on the worst skin pain HSSDD score is defined as a decrease from Baseline in HSSDD weekly worst skin pain score at or beyond the threshold for clinically meaningful change at Week 16. In this SAP, this endpoint is defined to include the exact value of the clinically meaningful threshold of 3, so that the skin pain response based on the HSSDD worst skin pain score is defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score at Week 16 among study participants with a score of ≥ 3 at Baseline.

The following other efficacy endpoints are included in the protocol but will not be included as part of the analysis:

- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HS Symptom Questionnaire (HSSQ) weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline
- Response on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48

The calculation of nominal p-values has been added for selected efficacy endpoints. These nominal p-values are not controlled for multiplicity and should not be used to declare statistical significance.

The protocol defines the PK-PPS separately by period, but there will only be one PK-PPS for the overall study.

4 STATISTICAL/ANALYTICAL ISSUES

4.1 Adjustments for covariates

The primary efficacy analyses and selected secondary analyses will be adjusted for the 2 randomization stratification variables:

- Hurley Stage at Baseline (II or III)
- Baseline antibiotic use (Yes or No)

If a participant is stratified in the incorrect stratum (ie, the stratum recorded in the Interactive voice or web Response System differs from the actual stratum the participant belongs to), the actual stratum will be used for the analysis.

The continuous secondary endpoints will also include the Baseline value as a covariate.

The Worst Skin Pain secondary endpoints (change from Baseline continuous endpoint and pain response binary endpoint) will also include analgesic use as a covariate.

4.2 Handling of dropouts or missing data

4.2.1 Efficacy data

Different approaches will be used to handle missing data including how intercurrent events (defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to the given visit) will be considered. A composite strategy will be implemented in which a positive clinical outcome is defined as the study participant achieving HiSCR₅₀ at the given visit and not receiving systemic antibiotic rescue medication, and not discontinuing study treatment due to an AE or lack of efficacy through that visit.

4.2.1.1 Handling missing data for the primary efficacy endpoint

If study participants have an intercurrent event as defined in Section 3.9, then the primary efficacy variable at that timepoint and all subsequent timepoints (whether the data were observed or not) will be set to “nonresponse” as the study participant has not met the criteria for response based on the composite estimand defined in Section 8.2. All remaining missing data for the endpoint will be imputed using multiple imputation Markov-Chain Monte Carlo method (MI-MCMC)/monotone regression for the primary analysis.

In addition, sensitivity analyses using NRI, MI-MCMC/reference-based methods, tipping point analysis, and observed case (OC) methods will be performed, which will assess the impact of different methods of handling missing data.

4.2.1.2 Handling missing data for the secondary efficacy endpoints

For secondary binary efficacy endpoints, intercurrent events will be handled, and missing data will be imputed, using the same methods as for the primary efficacy endpoint. NRI and OC methods will be performed as sensitivity analyses.

For secondary continuous efficacy endpoints, MI-MCMC/monotone regression is the primary method for imputing missing data, regardless of whether the missing data are preceded by an intercurrent event. That is, if an intercurrent event occurs on or before a visit, the result for that visit will be treated as missing and then imputed. If the imputation model cannot converge, last observation carried forward (LOCF) will be used. The OC method will be performed as a sensitivity analysis.

4.2.1.3 Handling missing data for the other efficacy endpoints

For other binary efficacy endpoints, missing data will be imputed using the same method as the primary efficacy endpoint. NRI and OC methods will be performed as sensitivity analyses of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀.

For other continuous efficacy endpoints, the MI-MCMC/monotone regression method will be used to impute missing data as the primary method, regardless of whether the missing data are preceded by an intercurrent event. That is, if an intercurrent event occurs on or before a visit, the result for that visit will be treated as missing and imputed with the missing data. If the imputation model cannot converge, LOCF will be used.

For other ordinal endpoints (EQ-5D-3L, PGI-S-HS, PGI-C-HS, PGI-S-SP, PGI-C-SP), the OC method will be applied as the primary analysis method. No imputation is applied.

4.2.1.4 Missing Data Overview and Summary

In summary, the approaches listed below will be used in this study for handling missing data for efficacy endpoints as appropriate:

- **NRI:** Participants who have missing data at the timepoint of interest are treated as though they did not respond to the treatment. This approach is also referred to as Composite Estimand (NRI).
- **Multiple Imputation (MI) – MCMC / Monotone Regression:** Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using monotone regression.
- **MI-MCMC / Reference-based imputation:** Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using an imputation model based on placebo (reference) data.
- **LOCF:** Post-Baseline missing data are imputed by carrying forward the last available observation (including Baseline).
- **Tipping point analyses:** Assumptions will be made about average outcomes among the subsets of participants who prematurely discontinued study treatment. Various “delta adjustments” will be made to the assumed responses among missing data in each treatment arm with varying degrees of plausibility in order to identify assumptions about the missing

data under which the conclusions change (O’Kelly, 2014). Then, the plausibility of such assumptions is discussed.

- **Observed case (OC):** Missing data are not imputed. Only participants with available data who have not discontinued study treatment at the given timepoint are considered. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing.
- **Treatment policy strategy:** All available data observed at the time point of interest will be considered, regardless of the occurrence of intercurrent events. This means the analysis includes on- and off-treatment values collected after study participants prematurely discontinued study treatment but agreed to remain on the study and continued to attend visits and provide assessments at those visits as well as values from study participants who received rescue antibiotic medication. Those observed values will be analyzed according to the study participant’s randomized treatment. Study participants for whom efficacy data cannot be obtained at the week of interest, despite attempts to retain them in the study, will have their data imputed using MI – MCMC / monotone.

The following table depicts which missing data handling approaches will be used based on endpoint priority (primary, secondary, other) and endpoint type (responder, continuous, ordinal).

Table 4–1: Missing data handling approach by endpoint priority and type

End-point Priority	Endpoint Type	Composite Estimand (NRI)	Modified Composite Estimand (MI)	MI (MCMC/Reference-based)	Tipping Point	Treatment Policy	Hypothetical Estimand	OC
Primary	Responder	S ^a	P	S ^a	S	S ^a		S
Secondary included in the statistical testing procedure	Responder	S ^a	P					S
	Continuous						P	S
Secondary not included in statistical testing procedure	Binary	X	X					X
	Continuous						X	X
Other	Responder	X ^d	X					X ^d
	Continuous						X	X ^b
	Ordinal						X ^c	X ^c

MI=multiple imputation, NRI=Nonresponder imputation, OC=Observed case, P=Primary method, S=Sensitivity method, X=Method to be used (no priority designated).

Note: Composite estimand (NRI) refers to the approach in which data preceded by the intercurrent event of study treatment discontinuation due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are also imputed as nonresponse.

Note: Modified Composite Estimand (MI) refers to the approach in which data preceded by the intercurrent event of study treatment discontinuation due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are imputed via a multiple imputation model.

Note: Hypothetical Estimand (MI) refers to the approach where outcomes for study participants without an intercurrent event of study treatment discontinuation are as observed, and outcomes for study participants with the intercurrent event are imputed via a multiple imputation model.

^a Imputation method is applied on continuous data, and responder endpoint is derived from the continuous endpoint based on complete data set where applicable.

^b Required only for by-visit summaries of variables whose value at Week 16 is part of the hierarchical testing procedure.

^c For variables with multiple categories, data will be summarized as observed with an additional missing row to capture missing data at a given visit.

^d NRI/OC sensitivity analysis will be performed only for HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀ summaries.

4.2.2 Missing data algorithms for efficacy analyses

These descriptions focus on the missing data procedures themselves and do not specifically account for dealing with intercurrent events, which is addressed in their respective sections.

4.2.2.1 MI – MCMC / Monotone Regression

In many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied.

Binary endpoint

For a binary endpoint (eg, HiSCR₅₀), the procedure is as follows:

1. Create a data set, one for each treatment group of participants with observed values and those needing estimation by multiple imputation. For the imputation step, a distinction is made between non-monotone missing values (ie, intermittent missing values between completed assessments) and monotone missing values (ie, where all participants data are missing after a given time point).
 - a. For the intermittent missing values, the missing values in each data set will be filled in using the MCMC method with multiple chain, monotone missing data imputing pattern, and non-informative prior for all parameters. Unless specified differently, the first 200 iterations will not be used (the “Burn-in” option). A total of 100 sets of imputations will be performed. The seed used for these imputations will be 762 and all other multiple imputation procedures described in this SAP will use this same seed as well. The resulting 100 imputed data sets will have a monotone missing pattern and will be imputed using a method for monotone missingness. Note that a separate imputation procedure must be invoked in SAS for each treatment group as the seed cannot be set for by groups beyond the first when using a by statement in PROC MI.
 - b. For monotone missing data, monotone regression will be used to impute missing data. A separate regression model is estimated for each variable with missing values (ie,

measurement at each time point). Based on the resulting model, a new regression model is then drawn and is used to impute the missing values for the variable. Since the data set has a monotone missing data pattern, the process is repeated sequentially for variables with missing values. This procedure will be based on the 100 imputed datasets generated from the MCMC procedure and will be performed by imputation. The SAS® PROC MI procedure will be used for the imputation.

In both cases, Hurley Stage at Baseline, Baseline antibiotic use, and value of the variable of interest at Baseline and at each post-Baseline visit (prior to the time point of interest) will be included in the imputation model. The post-Baseline values will need to be specified in chronological order in the imputation model so that SAS® PROC MI imputes variables from left to right (eg, the Week 2 value will be first imputed using regression based on the Baseline value, and then Week 4 value will be imputed using regression based on Baseline and Week 2 values, etc). The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model based on the MCMC method will only allow joint multivariate normal variables. Therefore, Hurley Stage at Baseline and Baseline antibiotic use will be re-coded as indicator variables. For Baseline antibiotic use, this will simply be 0 for Baseline antibiotic non-users and 1 for Baseline antibiotic users. For Hurley Stage at Baseline, this will be 0 for Hurley Stage II participants and 1 for Hurley Stage III participants. In order to achieve model convergence, Baseline antibiotic use may be dropped from the model. If convergence is still not obtained, then Hurley Stage at Baseline may also be dropped from the model. Additionally, if a variable is dropped in order to allow convergence for one model in a study, that variable does not have to be dropped from other models in the study if the model converges without dropping the variable. In other words, model convergence should be evaluated for each efficacy variable independently.

Note: The imputation of each lesion type (inflammatory nodule, abscess, draining tunnel, etc) will be performed separately. The 100 data sets obtained for each type will be merged by imputation number and subject number.

2. For each complete imputed data set, the dichotomous responder variable (eg, HiSCR 0 or 1) will be computed. Each complete imputed data set will then be analyzed based on the logistic regression model.

Note: For derivation of HiSCR response, the AN, inflammatory nodule, abscess, and draining tunnel (fistula/sinus tract) counts at Week 16 in the imputed data sets will be compared directly to the observed Baseline counts to determine response. If values outside of the pre-defined range of values for lesion count (<0) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed draining tunnel (fistula/sinus tract) count of -1 would be changed to 0 before deriving the HiSCR responder variable. Additional ranges for values for secondary and other endpoints are defined in Table 4-2.

Note: Standard rounding rules will also be applied to the imputed values of endpoints that can only take integer values (eg, abscess count). For example, if a study participant has an abscess count imputed as 2.4, this imputed value would be rounded down to 2. This rounding step is performed after the multiple imputation but before deriving the responder variable.

Table 4–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
Lesion count ^a	0	--	Yes
DLQI total score	0	30	Yes
hs-CRP	LLOQ/2	--	No
HSSDD item score	0	10	No
HSSQ item score	0	10	Yes
HiSQOL symptom status score	0	16	Yes
HiSQOL psychosocial impact score	0	20	Yes
HiSQOL impact on physical activities score	0	32	Yes
EQ-5D-3L VAS	0	100	Yes
WPAI dimension scores	0	100	No for variables: “Percent work time missed due to problem” and “Percent overall work impairment due to problem”. Yes for variables: “Percent impairment while working due to problem” and “Percent activity impairment due to problem”. These two variables can only take values that are multiples of 10.

^a Lesion counts will be imputed separately for each lesion type (abscesses, draining tunnels [fistulas/sinus tracts], inflammatory nodules, non-draining tunnels [fistulas/sinus tracts], non-inflammatory nodules, HS scars). The imputed lesion counts will be used to derive the endpoints that are dependent on the lesion count data (eg, HiSCR₅₀).

- Estimates of the adjusted responder rate for each treatment group and the associated SE are obtained from the logistic regression of each of the 100 imputed data sets. These estimates will be combined for overall inference using Rubin’s rules, which account for the uncertainty associated with the imputed values (Rubin, 1987), and the combined estimates and SEs will be used to construct 95% CIs using the logit scale. This will be done using SAS PROC MIANALYZE. The combined estimates and 95% CIs on the logit scale will be back-transformed using the inverse logit link function to obtain the adjusted responder rates (%) and associated 95% CIs.

Note: The (unadjusted) proportion of responders will be calculated at each time point by treatment group from the imputed datasets using SAS PROC FREQ. These results will also be combined into an overall inference using SAS PROC MIANALYZE.

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

As the estimates of the odds ratios from the logistic regression models in Step 3 follow a log-normal distribution, a log transformation is needed to normalize these 100 odds ratio estimates. That is because the procedures for combining results from multiple imputed datasets assume that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (Step 3). Additionally, the SE for the odds ratios are transformed as follows:

$$SE = \frac{\log(UCL) - \log(LCL)}{2Z_{\alpha/2}}$$

Where UCL and LCL are the upper and lower confidence limit, respectively, for the CI of the odds ratio from the logistic regression model, and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). The estimates of the log odds ratio for Bimekizumab relative to placebo and the corresponding upper and lower CLs will be provided. The odds ratio will be then estimated by exponentiating the estimate of the log odds ratio. The odds ratio and the confidence limits of the odds ratio will be estimated as follows:

$$OR = \exp(\text{Log odds ratio estimate})$$

$$LCL = OR * \exp(-SE * Z_{\alpha/2})$$

$$UCL = OR * \exp(SE * Z_{\alpha/2})$$

Where OR is the back-transformed estimate of the odds ratio just described, SE is the SE of the log odds ratio and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). These calculations will be done such that odds ratios and corresponding CIs are calculated for the odds ratio of bimekizumab vs. placebo.

Note: If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.

In addition to calculating the odds ratio, associated CIs, and p-values for the pairwise comparisons of bimekizumab and placebo, the estimated proportion of responders (ie, estimated responder rate) and the difference in the proportion of responders between each bimekizumab treatment group and placebo will be estimated, and 2-sided 95% CIs will be created for each difference. The creation of the estimates of the differences will be completed for each bimekizumab treatment group using the process detailed below:

1. Use the logistic regression model to calculate:

Least squares mean estimates of the log odds of bimekizumab (G_B) and placebo (G_P), as well as their corresponding standard errors (S_B and S_P , respectively).

Standard error of the least squares mean estimate of the log odds ratio (S_R)

2. Compute estimates for predicted proportions using the following transformations:

$$P_B = \exp(G_B) / (1 + \exp(G_B))$$

$$P_P = \exp(G_P)/(1 + \exp(G_P))$$

The difference in proportions is then given by:

$$D = P_B - P_P$$

3. Estimate the standard error of D by:

$$S_D = \sqrt{P_B^2(1-P_B)^2S_B^2 + P_P^2(1-P_P)^2S_P^2 + P_B(1-P_B)P_P(1-P_P)S_R^2 - P_B(1-P_B)P_P(1-P_P)(S_B^2 + S_P^2)}$$

The MCMC method for multiple imputation, as previously outlined, will be used to account for missing values. The calculation steps described above will be based on the results provided from the logistic regression model of the multiple imputed datasets. An estimate for the difference in proportions, D, and corresponding standard error, SD, will be computed for each of these datasets. The results from these analyses will be combined into a single estimate of the difference in predicted proportion of response and a 2-sided 95% CI interval using SAS PROC MIANALYZE.

Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. However, in the event there are computational challenges with the imputation model (eg, due to a standard deviation of 0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Continuous endpoint

For continuous endpoints (eg, Change from Baseline in DLQI total score at Week 16), the MI method will be applied as follows:

1. The MCMC/monotone regression method described above in Step 1 for binary endpoints will be performed.
2. Based on the multiply imputed data sets obtained for the given variable, the change from Baseline will be derived for each of the 100 complete imputed data sets based on the observed Baseline value and the observed/imputed post-Baseline values. Note that if the value itself is being summarized, no additional derivation is needed.
3. If a statistical model is being used for the analysis of the variable, then that will be performed for each imputation in this step. If no statistical model is being used, then simple descriptive statistics will be calculated.
4. For data excluding hs-CRP, the following rules apply. The results of the 100 imputed data sets (based on the statistical model or descriptive statistics) are combined with means and standard errors calculated using Rubin's rules (via PROC MIANALYZE). Note that for the calculation of other descriptive statistics such as the median, min, and max, Rubin's rules do not apply. MI estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. Thus, for median, Q1, Q3, minimum and maximum, the following approach will apply:

- The data will be summarized by treatment, visit, and imputation, and the summary statistics will be computed.
- Results will be summarized by taking the mean value of each summary statistic at each visit across all imputations.
- The number of decimal places will remain the same as the original for display purposes (ie, if the mean was presented to 1 decimal place, then the mean of the means will also be presented to 1 decimal place).

For hs-CRP only, the following rules apply. The hs-CRP data will be presented using the geometric mean, 95% CI for the geometric mean, median, Q1, Q3, minimum and maximum. The change from Baseline will be expressed as the ratio to Baseline in the summaries. The following approach will be applied:

- Following the MI procedure, the ratio to Baseline will be calculated for any of the imputed values
- The natural logarithm of the absolute values and of the ratios to Baseline will be calculated
- The logged values will be summarized (using PROC MEANS) by treatment, visit and imputation
- The datasets will be combined using PROC MIANALYZE in order to get the mean and 95% CI estimates from the absolute values and ratios to Baseline (based on logged data) across imputations
- The estimates of the mean and 95% CI will be back-transformed to obtain the geometric mean and 95% CI on the original scale
- For the median, Q1, Q3, minimum and maximum the procedure outlined above for the other endpoints will be followed

If the imputation model cannot converge, LOCF will be used.

4.2.2.2 MI – MCMC / Reference-based imputation

MI-MCMC / Reference-based imputation will be implemented as a supportive analysis for the primary efficacy endpoint (through Week 16).

In this case, placebo will be described as the reference arm.

This procedure will use an imputation model based on data from the placebo group only (Mallinckrodt, 2013). Reference-based MI assumes that the statistical behavior of the bimekizumab and placebo-treated participants after discontinuing study medication becomes that of the placebo-treated participants. All timepoints after discontinuation of the double-blind study treatment for both the bimekizumab and placebo groups will be considered missing. Multiple imputations are used to replace missing outcomes for bimekizumab- and placebo-treated participants who discontinued using multiple draws from the posterior predictive distribution estimated from the placebo arm. For binary efficacy endpoints (eg, HiSCR₅₀ at Week 16), imputation will be done on the lesion counts before assessing the imputed results for HiSCR₅₀ response.

The steps for the procedure are as follows:

1. For non-monotone (intermittent) missing data, MCMC will be used to impute lesion count data, with Baseline antibiotic use, Hurley Stage at Baseline, and lesion count at Baseline and at each post-Baseline visit (in chronological order) being included in the imputation model. This will be done only once for each participant in order to provide a dataset with monotone missing data.
2. Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcome data at visits $t=1, \dots, T$, where T is Week 16 for HiSCR₅₀.
 - a. *Initialization.* Set $t=1$ (Baseline visit)
 - b. *Iteration.* Set $t=t+1$. Create a data set combining records from bimekizumab- and placebo-treated participants with columns for covariates (Hurley Stage at Baseline and Baseline antibiotic use) and outcomes at visits 1 to t . Outcomes for all bimekizumab-treated participants are set to missing at visit t and set to observed or previously imputed values at visits 1 to $t-1$. Outcomes for placebo-treated participants are set to observed at visit t or observed or previously imputed values at visits 1 to $t-1$. The outcomes should be sorted in chronological order in the model.
 - c. *Imputation.* Impute missing values for visit t using previous outcomes for visits 1 to $t-1$, Baseline antibiotic use, and Hurley Stage at Baseline. Note that only placebo data will be used to estimate the imputation model since no outcome is available for bimekizumab-treated participants at visit t . Consequently, the input dataset should include all study participants from placebo but only study participants from the bimekizumab arm that have values at timepoint t missing.
 - d. Repeat steps 2a-2c, 100 times with different seed values (seeds ranging from 853 to 952) to create 100 imputed complete data sets. Study participants whose missing values were imputed in the last PROC MI call will be included in the input dataset for the next PROC MI call. Standard rounding rules will be applied to the imputed values. If the MI procedure yields a value outside of the pre-defined range for the given variable, the value will be updated to be within the predefined range of values for the endpoint of interest. For example, the imputed value for inflammatory nodules will be updated to 0 in the case of an imputed value less than 0.
 - e. *Analysis.* For each completed data set, the outcome (response or change from Baseline) will be calculated using the complete datasets (with no missing data).
3. Each complete imputed data set will then be analyzed based on the statistical model specified in this study (logistic regression). The Week 16 results from logistic regression of each of the 100 imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

4.2.2.3 Tipping Point Analysis

Tipping point analyses will also be implemented as a supportive analysis for the primary efficacy endpoint.

The objective of the tipping point analyses is to evaluate the sensitivity of results to departures from the missing at random assumption and to identify the point at which departures cause results to "tip" from statistically significant to statistically non-significant. As such, these tipping

point analyses will only be performed if the primary efficacy analysis results in a statistically significant treatment effect.

For tipping point analyses, data for participants after the intercurrent event date (See Section 3.9) will be changed to missing prior to imputation but will not be changed to non-response after imputation.

The worst-case scenario will be evaluated first. All missing primary endpoint values for study participants randomized to bimekizumab (where missing values include observations after the intercurrent event date and any other missing values) will be imputed as non-responders, while all missing values for placebo-randomized study participants will be imputed as responders. While there is little justification for such an approach, it makes the most putative assumption possible against a bimekizumab treatment effect. After applying this imputation approach, a logistic regression model consistent with the one described for the primary analysis will be applied. If the p-value for the odds ratio of bimekizumab versus placebo remains significant, then no further tipping point analyses are needed.

If this analysis based on the worst-case scenario results in a p-value that is not significant (eg, greater than 0.025), then additional tipping point analyses will be performed to identify the point at which results switch or “tip” from significant to non-significant. Note that each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo independently for these analyses. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the p-value in this analysis method will be 0.05 instead of 0.025 throughout for that dose. In the tipping point analysis, a shift parameter or delta adjustment is applied to missing, and subsequently imputed primary endpoint values (where missing values include observations after the intercurrent event and any other missing values). These delta implemented on the primary endpoint as follows:

1. Data after intercurrent event date (See Section 3.9) will be set to missing.
2. The same MCMC method described in Section 4.2.2.1 (Step 1a) will be implemented for non-monotone (intermittent) missing pattern values, using the same imputation model. This will be based on 100 sets of imputations.
3. Based on the 100 datasets obtained in Step 2, a monotone regression model will be applied (using the same imputation model as in Step 2) as described in Section 4.2.2.1 (Step 1b). This will be based on 1 imputation.
4. Delta adjustments will be made to imputed lesion count values at Week 16, independently in each treatment group as described below.
5. Delta adjusted imputed values will be truncated so that they are within the range of allowable values for each component.
6. Following the delta adjustments for the lesion counts, HiSCR₅₀ will then be derived based on the delta-adjusted multiply imputed data sets obtained for each component.
7. Each of the 100 imputed datasets will then be analyzed using a logistic regression model with factors of treatment group, Baseline Hurley Stage, and Baseline antibiotic use.

8. The results obtained from the 100 logistic regression analyses in Step 7 will be combined for overall inference using Rubin's rules, and the results obtained for each shift parameter will be presented in a single table.
9. Steps 4 to 8 will be repeated so that, at each iteration, missing values are adjusted with a larger delta than at the previous iteration. The process will go on until the p-value for the odds ratio between bimekizumab and placebo is no longer statistically significant (eg, ≥ 0.025). The odds ratio, 97.5% CI (or 95% depending on the significance level being used for testing), and p-values obtained for each value of delta will be combined in one single table.

The delta adjustments result in study participants randomized to bimekizumab with missing data having a lower probability of response compared to study participants randomized to placebo with missing data. Since HiSCR₅₀ response is an endpoint for which high lesion counts are associated with a less favorable outcome:

- A positive adjustment is applied to the imputed value for study participants randomized to bimekizumab in order to increase the imputed value and decrease the likelihood of response.
- A negative adjustment is applied to the imputed value for study participants randomized to placebo in order to decrease the imputed value and increase the likelihood of response.

To start, imputed values within each lesion type, will be adjusted by the same value in each treatment arm. This adjustment will be 5% of the observed range within that lesion type. Depending on the results obtained, this adjustment will be multiplied for step 9 above (2 times, 3 times the initial adjustment) until the p-value is no longer statistically significant.

Additionally, study participants randomized to bimekizumab with an intercurrent event should be set to non-response, after applying the delta adjustment outlined in Step 6 above. This ensures study participants randomized to bimekizumab do not have a higher probability of response in the tipping point analyses compared to the primary analysis (ie, a study participant randomized to bimekizumab who is non-responder in the primary analysis cannot become a responder in the tipping point analyses).

4.2.3 Rationale for estimand

Intercurrent events have been identified within the estimands for this study because of their potential to impact efficacy assessments linked with the primary and secondary study objectives. In order to account for the effect of any observed post-randomization intercurrent events on the efficacy analyses, the following estimand strategies will be implemented when evaluating the primary and secondary efficacy endpoints:

- A composite estimand strategy will be used for the primary analysis of the primary and binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, flare, HS worst skin pain response);
- A hypothetical estimand will be used for the primary analysis of the continuous secondary endpoints (CFB in DLQI total score and in "worst skin pain" item for the HSSDD).

4.2.3.1 Composite estimand

A composite estimand strategy as defined in Section 8.2.2 allows incorporation of the two intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) within the definition of the endpoint. These

intercurrent events are considered meaningful to the efficacy outcome following receipt of study medication. For example, within the proposed composite estimand framework, a randomized study participant who discontinues from study treatment due to lack of efficacy prior to Week 16 will be considered a treatment failure at Week 16 regardless of the lesion count assessment performed at that visit.

The assumptions and robustness of the primary analysis (modified composite estimand as defined in Section 8.2.2) will be assessed through the sensitivity analyses defined in Section 8.2.3. The impact of intercurrent event handling and data imputation methods on endpoint derivation will also be assessed via the analyses of lesion counts and derived HiSCR variables as specified in Section 8.4.2.1 and Section 8.4.1.1, respectively.

4.2.3.2 Hypothetical estimand

The hypothetical estimand is defined in Section 8.3 and involves a data-driven approach to account for the potential impact of intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) on the analysis of continuous efficacy endpoints. Under this framework, outcomes for study participants without an intercurrent event are analyzed as observed. Conversely, outcomes for study participants with an intercurrent event are imputed via a multiple imputation model, ie any recorded data on or after the intercurrent event will be set to missing and imputed via multiple imputation following the strategy established in Section 4.2.2.1.

4.2.4 Dates and times

For analyses of AEs and concomitant medication usage, a complete date is required in order to correctly identify the AE or medication as occurring during treatment or not, and for correctly assigning an AE or concomitant medication to the Initial Treatment Period or Maintenance Treatment Period.

For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the participant data listings (ie, no imputed values will be displayed in data listings).

Partial AE and concomitant medication start dates will be imputed as follows:

- Imputation of Partial Start Dates
 - If only the month and year are specified:
 - If the month and year of first dose of study medication is the same as the month and year of the partial start date, then use the date of first dose of study medication,
 - Else, if the month and year of the partial start date are the same as the month and year of a study medication switch date, then use the date of study medication switch,
 - Otherwise, use the 1st of the month of the partial start date;
 - If only the year is specified:
 - If the year of first dose of study medication is the same as the year of the partial start date, then use the date of first dose of study medication,

- Else, if the year of the partial date is the same as the year of a study medication switch date, then use the date of study medication switch,
- Otherwise, use the 1st of January of the year of the partial start date;
- If the start date is completely unknown:
 - If the stop date is unknown or not prior to the date of first dose of study medication, then use the date of first dose of study medication,
 - If the stop date is prior to the date of first dose of study medication, then use the 1st of January of the year of the stop date.
- Imputation of Partial Stop Dates
 - If only the month and year are specified, :
 - Use the last day of the month of the partial stop date;
 - If only the year is specified
 - use December 31st of the year of the partial stop date;
 - If the stop date is completely unknown,
 - Do not impute the stop date.

Note that if the stop date or the imputed stop date is prior to the imputed start date, then follow the procedure outlined below:

- If only the year of the start date is specified:
 - If the year of start date is the same as the year of first dose of study medication and the imputed stop date is after the date of first dose of study medication, then set the start date to the date of first dose of study medication,
 - Otherwise, set the 1st January of the year of the start date;
- If only the month and year of start date are specified:
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is on or after the date of first dose of study medication then set the start date to the date of first dose of study medication,
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is before the date of first dose of study medication then set the start date to the 1st of the month of partial start date.

Missing start times for medications will be imputed as 00:00h or with the time of dosing for events occurring on the date of IMP administration in case of missing hour and minute. Otherwise start times with only missing minutes will be imputed with :00 or with the minutes of dosing for events occurring on the date and hour of IMP administration.

In the event of ambiguity or incomplete data that makes it impossible to determine whether a medication was concomitant or an AE was treatment emergent, the medication will be considered as concomitant or the AE will be considered treatment emergent. Similarly, in the event of ambiguity or incomplete data which makes it impossible to determine whether a

medication or AE is to be assigned to the Initial Treatment Period or to the Maintenance Treatment Period (or both, for medications), then the medication will be assigned to both Treatment Periods, and the AE will be assigned to the Initial Treatment Period.

4.3 Interim analyses and data monitoring

4.3.1 Data monitoring committee

An independent data monitoring committee (DMC) will periodically review unblinded efficacy and safety data to assess the benefit/risk of bimekizumab in study participants with moderate to severe HS. Efficacy data summaries and individual study participant-level data listings may be provided to the DMC to put the safety review in the context of risk/benefit. Any data to be provided is specified per the DMC charter. [REDACTED]

4.3.2 Interim analysis

4.4 Multicenter studies

The center-by-treatment interaction will be tested by adding center and a center-by-treatment interaction term (Section 8.2.3.11). In the model, center will be based on the original centers prior to pooling (Section 3.7). However, if the model is unable to converge due to a low number of participants at a given center, a pooling by center will be applied in order to allow the model to converge. If convergence is still not achieved, a pooling by region will be applied. If convergence still cannot be achieved, this analysis will not be performed. Detailed strategy in Section 3.7 will be applied.

4.5 Multiple comparisons/multiplicity

To control the overall type I error rate at 0.05 for the multiple comparisons in the primary and secondary efficacy endpoints, a closed testing procedure under a parallel gatekeeping framework will be applied (Sun, 2018).

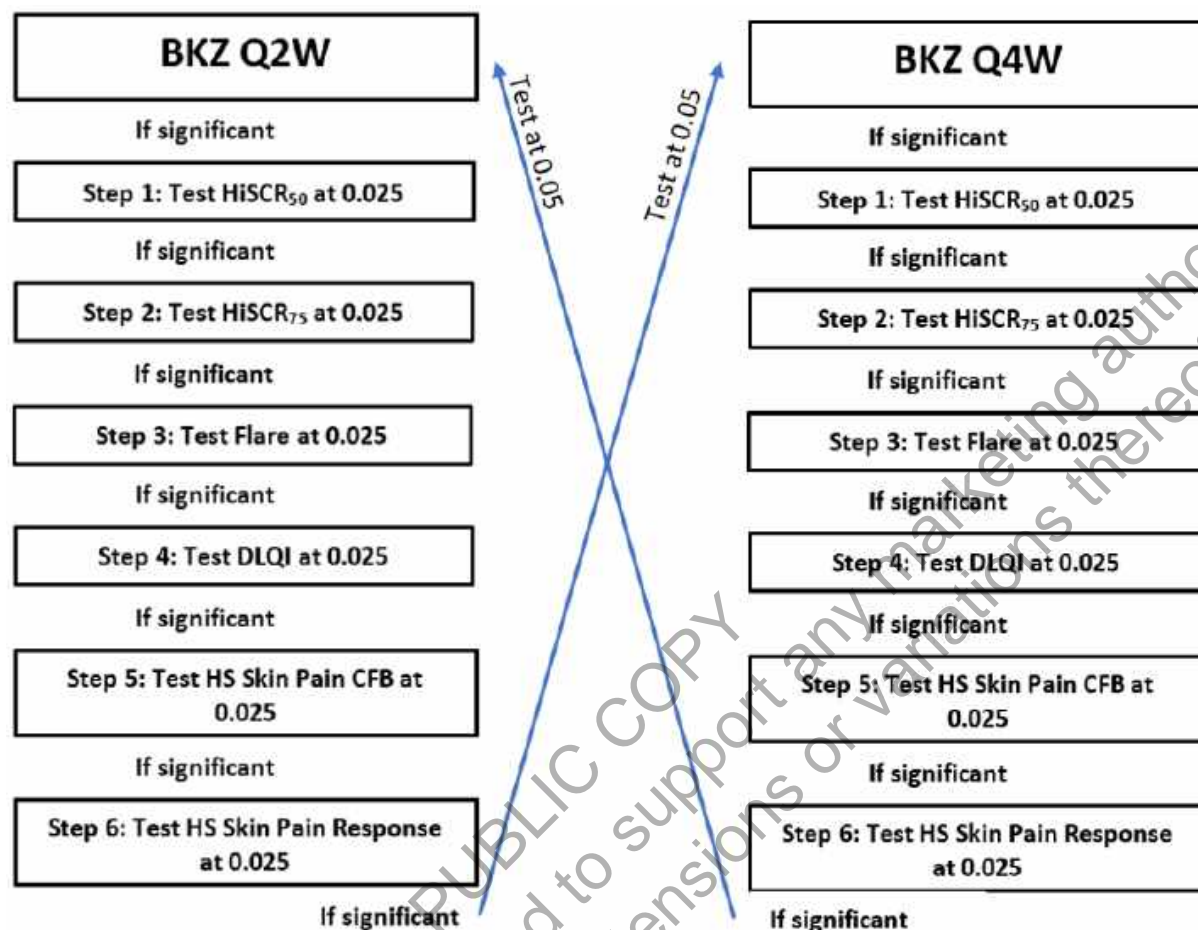
Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$). Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

1. Step 1: Test HiSCR₅₀ at significance level 0.025.
2. Steps 2 to 6 – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown in Figure 4-1, moving to the next step only if significance achieved at 0.025.
3. In the event that Step 6 is significant at 0.025 for a given dose, then Steps 1 to 6 will be repeated for the other dose using a significance level of 0.05.

The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment using the analysis methods specified in Section 8.3:

1. Proportion of study participants who achieve HiSCR₇₅ at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
2. Proportion of study participants who experience at least 1 flare by Week 16, with flare defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
3. Absolute CFB in DLQI Total Score at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
4. Absolute CFB in Skin Pain Score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HSSDD.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
5. Skin pain response at Week 16, based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo

Figure 4-1: Sequence of testing



AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks
HS skin pain response is tested among study participants with a score of ≥ 3 at Baseline.

4.6 Use of an efficacy subset of participants

A sensitivity analysis of the primary endpoint will be performed based on the FAS, the PPS, and the CFS.

4.7 Active-control studies intended to show equivalence

Not applicable.

4.8 Examination of subgroups

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, flare, and worst skin pain response endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period. Additional subgroup analyses will be performed on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, flare, and skin pain response endpoints which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

The following subgroup variables will be determined using Baseline data, except for analgesic use, lesion intervention, and antibody positivity:

- Age (<40 years, 40 to <65 years, ≥65 years)
- Gender (male, female)
- Disease duration (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for analysis.

- Region (North America [Canada, USA], Western Europe [France, Germany, Ireland, Italy, Spain, United Kingdom], Central/Eastern Europe [Bulgaria, Czech Republic, Hungary, Poland], Asia/Australia [Australia, Israel, Japan])
- Weight (≤100 kg, >100 kg)
- BMI (<25 kg/m², 25 to <30 kg/m², ≥30 kg/m²)
- Race (Black or African American, White, All Other Races [American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other/Mixed])
- Systemic antibiotic therapy at randomization (yes, no)
- Prior biologic therapy for any indication (yes, no)
- Prior biologic therapy for HS (yes, no)
- Hurley Stage at Baseline (II or III)
- Analgesic users (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period (Section 6.4.2 specifies how participants are classified as analgesic users)
- Lesion intervention (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period
- Antibody positivity (confirmatory assay: negative or positive; see Section 9.3.2)
- Antihistamine users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users) (applicable only to the skin pain response endpoint)

The following subgroups for analysis on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score will be determined based on medication use during the Initial Treatment Period:

- Antihistamines users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users)

- Analgesics users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as analgesic users)
- Systemic antibiotic therapy start/increase after randomization during the Initial Treatment Period (yes, no)

All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

5 STUDY POPULATION CHARACTERISTICS

5.1 Study participant disposition

Summaries of reasons for screen failures (for ES), disposition of participants (for ES), disposition of analysis sets (for RS), disposition and discontinuation reasons in the Initial Treatment Period (for RS) and the Maintenance Treatment Period (for MS), as well as the participants who discontinued due to AEs in the Initial Treatment Period (for RS) and Maintenance Treatment Period (for MS) will be produced. The disposition of participants for all participants screened will include the number of participants included in each analysis set (ES, RS, SS, FAS, AMS, MS, PPS, and PK-PPS) overall and by site.

Participants are defined as completing the Initial Treatment Period if they have a Week 16 visit, or if they fail to attend the Week 16 visit but attend at least one visit in the Maintenance Treatment Period.

The following listings for participant disposition will be provided: participants who did not meet study eligibility criteria (for ES), participant disposition (for ES), participant discontinuation (RS), visit dates (for RS), participant analysis sets (for ES), participants excluded from efficacy analysis (for RS).

To assess participant disposition (entry and periods in the study) during the COVID-19 pandemic, study participant disposition will also be assessed by period of the COVID-19 pandemic (pre – during – post), by comparing the dates of visits (or events) to the dates of the COVID-19 pandemic period. The dates to categorize the periods of the COVID-19 pandemic (pre/during/post) are defined below:

- Pre-COVID-19 pandemic period: Period prior to COVID-19 pandemic start date defined as 11-Mar-2020
- COVID-19 pandemic period: Period from 11-Mar-2020 through the COVID-19 pandemic end date which is currently not defined at the time of approval of the SAP
- Post-COVID-19 pandemic period: Period after the declaration of the end of the pandemic

5.2 Impact of COVID-19

A listing of visits affected by COVID-19 will be presented based on the ES including the visit, date of visit, relationship to COVID-19, impact category and a narrative (short description) of the event. These data will be summarized for non-randomized participants and by treatment group and overall, for enrolled participants.

A summary of study visits by COVID-19 pandemic period (pre – during – post) will be presented for participants enrolled prior to and during the pandemic.

In addition, in order to assess the potential impact of COVID-19 on the collection and reporting of efficacy data, a separate summary on the RS will be presented to display missing data as well as data collected via an alternative modality (e.g.: phone, video call) for efficacy endpoints included in the hierarchy (Section 4.5). For these displays, missing data will be presented only for visits affected by COVID-19, as reported on the dedicated eCRF page. Missing data at other visits and for other reasons will not be included. Note that the remote contingencies for COVID-19 or other exceptional circumstances are not applicable to efficacy assessments and documentation (eg, lesion-based assessments, photography) that require direct face-to-face physician/participant interaction.

5.3 Protocol deviations

Summaries, based on the RS and the MS, displaying the number and percentage of participants with an important protocol deviation (including a summary of participants excluded from the PPS or PK-PPS due to important protocol deviations) by treatment group in the Initial Treatment Period and in the Maintenance Treatment Period, respectively, will be provided. A separate summary of participants with protocol deviations related to COVID-19 will be provided.

A by-participant listing of protocol deviations will be provided. Protocol deviations (eg, missing assessments or visits) related to COVID-19 will be listed separately.

6 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

All summaries detailed in this section will be performed on the RS by treatment group. Summaries for demographics and other baseline characteristics will also be repeated on SS and MS. If the RS and SS are identical, the SS summaries will not be created.

6.1 Demographics

Demographic variables will be summarized by treatment group and overall.

The following continuous variables will be summarized using descriptive statistics (number of study participants, mean, SD, minimum, median, and maximum).

- Age (years)
- Height (cm)
- Weight (kg)
- BMI (kg/m²)

BMI (kg/m²) will be calculated as:

$$\text{BMI} = \frac{\text{Weight (kg)}}{(\text{Height (m)})^2}.$$

The following categorical variables will be summarized generally using frequency counts and percentages.

- Age group (≤ 18 , $19 < 65$, ≥ 65 years)
- Age group ($18 < 65$, $65 < 85$, ≥ 85 years)

- Age group (<40, 40-<65, ≥65 years)
- Body weight (≤100 kg, >100 kg)
- Gender
- Race
- Ethnicity
- Ethnic subgroup
- BMI (<25 kg/m², 25 to <30 kg/m², ≥30 kg/m²)
- Region
- Smoking history
- Country

By-participant listing of demographics for all study participants screened will be provided.

Childbearing potential and lifestyle will be collected at Screening.

6.2 Other Baseline characteristics

The following Baseline disease characteristics will be summarized by treatment group:

- Lesion counts by anatomical region and lesion type, total lesion counts across anatomical regions by lesion type, Hurley Stage by anatomical region and worst overall Hurley Stage across anatomical regions
- IHS4 score, individual items of the HSSDD, HS-Physician's Global Assessment, DLQI total score and HiSQOL domain and total scores
- hs-CRP
- Duration of disease (years)

Duration of disease (years) will be calculated as:

$$\text{Disease Duration} = \frac{(\text{Date of randomization} - \text{Date of HS Diagnosis}^1)}{365.25}$$

¹If the date of HS diagnosis is partial, it will be imputed to the most recent feasible date (ie, last day of the month if only day is missing, or the last day of the year if day and month are missing). Note that if the date of randomization is missing then the duration of disease will be derived using the date of screening. If the imputed date results in a participant having a disease duration of less than 6 months and the inclusion criterion related to having HS for at least 6 months is confirmed to not have been violated, then the participant's duration of disease will be set to 6 months. If that criterion has been violated, then the participant's duration of disease will be the imputed value of less than 6 months.

- Duration of disease (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for the summary.

- Baseline antibiotic use (yes, no) (According to the randomization strata)
- Baseline antibiotic use (yes, no) (Derived)
- Hurley Stage at Baseline (According to the randomization strata)
- Hurley Stage at Baseline (Derived)

In addition, the following Baseline disease characteristics will be summarized by the derived Baseline Hurley Stage and by the derived Baseline antibiotic use and treatment group for the RS:

- IHS4 score
- “worst skin pain score” and “average skin pain score” in the HSSDD
- HS-Physician’s Global Assessment
- DLQI total score
- hs-CRP
- Duration of disease (years)
- Total lesion counts

6.3 Medical history and concomitant diseases

Previous and ongoing medical history will be summarized by treatment groups, system organ class (SOC), and preferred term (PT) using MedDRA[®]. Medical procedures are not coded.

The following listings for medical history and concomitant diseases will be provided: medical history, HS history, concomitant medical procedures, previous and ongoing medical history glossary, previous and ongoing medical history conditions, and procedure history.

6.4 Prior and concomitant medications

Prior medications include any medications that started before the start date of study medication. Concomitant medications are any medication that has a start date on or after the start date of study medication, or any medication that has a start date on or before the last dose of study medication + 28 days (whether placebo or bimekizumab).

Any medication that started before the first dose of IMP and continued after will be classified as both prior and concomitant. Such medications will therefore be counted in the tabulations for both prior and concomitant medication.

Details of imputation methods for missing or partial dates are described in Section 4.2.4.

The number and percentage of participants taking prior medications will be summarized by treatment group, overall and by ATC class, presenting Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT. Prior antibiotic medications will be summarized similarly.

The number and percentage of participants taking concomitant medications will be summarized similarly for the Initial Treatment Period and Maintenance Treatment separately. The number and percentage of participants taking concomitant antibiotic medications, antihistamines, and analgesics will be summarized separately by treatment group, overall, and by ATC class,

presenting Anatomical Main Group (ATC Level 1), Pharmacological Subgroup (ATC level 3), and PT for the Initial Treatment Period and Maintenance Treatment Period separately.

Separate summaries will be presented for participants taking rescue medication for the Initial Treatment Period and Maintenance Treatment Period separately, identified by a 'yes' response to the 'Is this a rescue medication' question on the electronic case report form (eCRF). This summary will be performed separately for analgesic use and antibiotic use.

Additional summaries for the Initial Treatment Period and Maintenance Treatment Period will be presented for participants taking systemic antibiotic medications that qualify as intercurrent events as described in Section 3.9.

The number and percentage of study participants with concomitant vaccines for COVID-19 will be summarized by treatment group, overall and by World Health Organization Drug Dictionary Standardized Drug Grouping (SDG), presenting SDG subgroup, and preferred term. The SDG subgroup Vaccines for COVID-19 will be used to identify vaccines for COVID-19 using the narrow scope; this subgroup is divided further into separate subgroups which is the level that will be presented. The number of individual occurrences of the vaccine for COVID-19 will also be summarized.

A listing of concomitant vaccines for COVID-19 will be provided.

6.4.1 Assignment of medications to study period

The following rules will be used to assign a concomitant medication to a study period:

- **Initial Treatment Period:** a medication will be assigned to the Initial Treatment Period if it has been taken at least once between the first administration of IMP on Day 1 up to Week 16. This includes medications that started prior to the Initial Treatment Period and those that continued into the Maintenance Treatment Period.
- **Maintenance Treatment Period:** a medication will be assigned to the Maintenance Treatment Period if it has been taken at least once between Week 16 and the final visit. This includes medications that started prior to the Maintenance Treatment Period.

Thus, a medication taken from the time of the first drug administration in the Initial Treatment Period to any timepoint after Week 16 will be assigned to both the Initial Treatment Period and the Maintenance Treatment Period.

Methods for dealing with partial dates are specified in Section 4.2.4.

6.4.2 Classification of participants as analgesic, antihistamine users

If a participant has taken a new analgesic/increased regimen of analgesic, or taken an antihistamine, on 1 or more days (need not be consecutive) in a study period (Initial Treatment Period or Maintenance Treatment Period), then for that period the participant will be classified as an analgesic or antihistamine user, respectively. The period under consideration is to match the period as defined for the HSSDD for the Initial Treatment Period or HSSQ for the Maintenance Treatment Period, based on dates/times of the medications taken.

New analgesic/increased regimen of analgesic use, regardless of indication, is defined as an analgesic medication with start date on or after the first dose of study medication. Stable

analgesics (ie, analgesics which were taken already before randomization) will not be included in this category of analgesic user. This classification will be used for selected subgroup analyses.

Antihistamine use is identified by considering the ATC classification. This classification is used for analyzing the Worst Itch endpoint and for selected subgroup analyses, by visit, for the Initial Treatment Period and Maintenance Period as applicable.

Additionally, if a participant has taken a new analgesic/increased regimen of analgesic on 1 or more days (need not be consecutive) prior to the Week 16 visit, then for that week the participant will be classified as an analgesic user. This classification will be used to adjust the formal analysis of the Worst Skin Pain secondary endpoints. If there is a visit date but no HSSDD available at the visit, then the analgesic/antihistamine user status for that week will be derived based on the visit date. If there is no visit available, then the weekly analgesic/antihistamine user status will default to the analgesic/antihistamine status for the overall study period.

7 MEASUREMENTS OF TREATMENT COMPLIANCE

Due to the method of administration of the treatments, compliance will be examined in terms of completed injections.

Treatment compliance will be calculated as:

$$\frac{N_{actual}}{N_{expected}} \times 100\%$$

where N_{actual} is the total number of actual (completed) injections, and $N_{expected}$ is the total number of expected injections. In this study, dosing occurs every 2 weeks from Week 0 to Week 46, where 2 injections are administered at each given visit either with active dosing or placebo injection. It is expected that a participant should complete a total of 48 injections by the end of study. If a participant discontinues early, then the number of expected injections is based on the time of early discontinuation relative to the dosing visits. For example, if a participant discontinues after Week 8 visit and prior to Week 10 visit, the total number of expected injections will be 10.

A summary of percent treatment compliance categorized as <75% and ≥75% will be provided by treatment group for each study period (Initial Treatment Period for the RS, Maintenance Treatment Period for the MS, and the combined Initial and Maintenance Treatment Period for the AMS).

A by-participant listing of treatment compliance will be provided.

8 EFFICACY ANALYSES

All efficacy analyses of primary, secondary, and other variables will be performed on the RS unless otherwise specified. All efficacy summary tables will be displayed by treatment sequence unless otherwise specified. The primary and secondary endpoints, and their components, will also be summarized by the derived Hurley Stage at Baseline (grouping each stage and overall) and treatment sequence and by the derived Baseline antibiotic use (yes/no and overall) and treatment group.

8.1 Lesion count assessment

The primary efficacy endpoint and some of the secondary and other efficacy endpoints discussed in Section 8.3 and Section 8.4 are based on the assessment and/or counts of different types of lesions in the following main anatomical regions at each visit.

- Inguinal (groin)
- Axillary (armpit)
- Chest/breast
- Gluteal
- Abdomen including supra pubic
- Back
- Head
- Neck
- Leg
- Other

These anatomical regions are further classified into the following locations, for the left and right sides of the body, as applicable:

- Inguinal excluding genital and pubic area
- Inguinal including genital and pubic area
- Submammary
- Intermammary
- Chest
- Breast
- Gluteal – Buttocks
- Gluteal – Perianal/Perineal
- Scalp
- Face

All “Other” anatomical regions with lesions present will be specified in free text on the eCRF.

The number of each of the following types of lesions will be recorded in each anatomical region, and then summed across all anatomical regions:

- Abscesses
- Inflammatory nodules
- Non-inflammatory nodules
- Draining tunnels (fistulas/sinus tracts)

- Non-draining tunnels (fistulas/sinus tracts)
- HS scars

If participants undergo lesion interventions as specified in the study protocol, the affected lesions will be counted by the Investigator as permanently present, thus accounting for potential bias due to the intervention. Section 8.4.23 specifies how the intervention data will be presented.

8.2 Primary efficacy endpoint

The primary and sensitivity analyses of HiSCR₅₀ response at Week 16 are summarized in Table 8–1.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

		Estimands for Primary Endpoint			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Primary Objective: To evaluate the efficacy of bimekizumab in study participants with moderate to severe HS					
HiSCR ₅₀	Primary (Section 8.2.2)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (nonresponders).	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.1)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.
HiSCR ₅₀	Sensitivity (Section 8.2.3.2)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – Reference-Based Regression under a missing not at random assumption.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.3)	HiSCR ₅₀ response at Week 16	RS	Composite strategy^a , as for the primary analysis.	A tipping point analysis will be used where various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. The odds ratio versus placebo is based on a logistic regression for each value of delta.
HiSCR ₅₀	Sensitivity (Section 8.2.3.4)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a treatment policy strategy , whereby the data from the Initial Treatment Period are used regardless of whether the intercurrent event occurred.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.5)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.
HiSCR ₅₀	Sensitivity (Section 8.2.3.6)	HiSCR ₅₀ response at Week 16	FAS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.7)	HiSCR ₅₀ response at Week 16	PPS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.8)	HiSCR ₅₀ response at Week 16	CFS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.9)	HiSCR ₅₀ response at Week 16	RS	The same two intercurrent events used for the primary analysis will be used. Any missing data due to COVID-19 will also be considered an intercurrent event. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (nonresponders).	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.10)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a stratified Cochran-Mantel-Haenszel (CMH) test. Missing values not preceded by an intercurrent event will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

CFS=Covid-19 Free Set; CMH=Cochran-Mantel-Haenszel; COVID-19=coronavirus disease 2019; FAS=Full Analysis Set; HiSCR=Hidradenitis Suppurativa Clinical Response; IES=intercurrent event(s) strategy; MCMC=Markov Chain Monte Carlo; MI= multiple imputation; PLS=Population-level summary; Pop=Population; PPS=Per-Protocol Set; RS=Randomized Set

^a The composite estimand strategy will be modified in the tipping point analysis such that participants with intercurrent events will be treated as nonresponders only in the bimekizumab treatment groups.

8.2.1 Derivation of HiSCR₅₀ at Week 16

The following algorithm will be applied to derive HiSCR₅₀ at each visit, based on total lesion counts across anatomical regions for the 3 relevant lesion types recorded as specified above in Section 8.1:

1. Calculate the AN count at each visit as the total number of abscesses plus the total number of inflammatory nodules, across all anatomical regions
2. Calculate the percentage change from Baseline in AN count (%ΔAN) at each visit as

$$100 \times (\text{AN count at post-Baseline visit minus Baseline AN count}) / (\text{Baseline AN count})$$

3. Calculate the change from Baseline in the abscess count by subtracting the Baseline abscess count from the abscess count at each post-Baseline visit
4. Calculate the change from Baseline in the draining tunnel (fistula/sinus tract) count by subtracting the Baseline draining tunnel (fistula/sinus tract) count from the draining tunnel (fistula/sinus tract) count at each post-Baseline visit
5. If the %ΔAN is less than or equal to -50%, and the change from Baseline in the abscess count is zero or negative (ie, non-positive) together with a non-positive change from Baseline in the draining tunnel (fistula/sinus tract) count, then the HiSCR₅₀ will be assigned a value of 1 (ie, HiSCR₅₀ is achieved); otherwise, the HiSCR₅₀ will be assigned a value of 0 (ie, HiSCR₅₀ is not achieved).

In cases where the inflammatory nodule, abscess or draining tunnel (fistula/sinus tract) count is missing and will not allow for the HiSCR₅₀ calculation, the rules for handling missing values in the analysis will be applied (Section 4.2.1 and Section 8.2.3).

The primary efficacy endpoint is attained if the participant has a HiSCR₅₀ of 1 at Week 16.

8.2.2 Primary analysis of the primary efficacy endpoint

The primary endpoint is the HiSCR₅₀ response at Week 16 and corresponding analyses are based on the RS. The primary efficacy analysis will evaluate the composite estimand in the RS as described in Table 8–1. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and completion of study treatment through Week 16 without receiving systemic antibiotic rescue medication or discontinuing IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for the primary efficacy analysis:

1. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
2. Study participant-level outcome=HiSCR₅₀ at Week 16.
3. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. More information is provided in Section 3.9. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, and not discontinuing study treatment due to an AE or lack of efficacy through Week 16. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation as defined in Section 4.2.1. The rationale for this composite estimand is provided in Section 4.2.3.1.

4. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

The statistical hypothesis for the HiSCR₅₀ response at Week 16 is that the conditional odds ratio for the HiSCR₅₀ response in the bimekizumab treatment group relative to the placebo group is equal to 1.

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, and Baseline antibiotic use. The odds ratio versus placebo, p-value (from Wald test), and 97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.

The number and percentage of participants who are HiSCR₅₀ responders at Week 16 will be summarized.

By-participant listings of HiSCR responder endpoints will be provided.

8.2.3 Sensitivity analyses of the primary efficacy endpoint

The following sensitivity analyses for the primary efficacy endpoint will be performed to evaluate the assumptions related to the handling of missing data. Details of the estimands for each analysis are described in Table 8-1.

8.2.3.1 Nonresponse imputation

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (ie, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) will be imputed as nonresponse. That is, participants who experience an intercurrent event will be imputed as nonresponder at the timepoint of the event and all subsequent timepoints (including any recorded data after the event), and all missing data will also be imputed as nonresponse.

The same analysis model as in the primary efficacy analyses will then be used on the imputed data set.

8.2.3.2 MI-MCMC / Reference-based imputation

Deviations from the missing at random pattern will be evaluated using a reference-based MI approach (see Section 4.2.2.2). Intermittent missing data will be imputed using the MCMC method. The remaining monotone missing data will be assumed to follow a missing not at random pattern. These data will be imputed using a reference-based approach in which the MI model is based on data from the placebo group, thereby assuming that monotone missing data follow a trajectory similar to the placebo group.

The same analysis model as in the primary efficacy analyses will then be used on the imputed data set.

8.2.3.3 Tipping point analysis

Tipping point analyses will be performed to evaluate missingness assumptions. Various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. It will include scenarios where study participants who have missing data and are randomized to bimekizumab have a

lower probability of response compared to study participants who have missing data and were randomized to placebo. This includes the worst-case scenario where study participants who have missing data and are randomized to bimekizumab are considered nonresponders, while study participants who have missing data and were randomized to placebo are considered responders. The goal of the tipping point analysis is to systematically vary assumptions until there is no evidence of a treatment effect (if possible). The plausibility of such required delta adjustment will then be assessed.

Refer to Section 4.2.2 for more details on the methodology.

8.2.3.4 Treatment policy

The treatment policy strategy for addressing intercurrent events will be considered. This will be based on an analysis of all available data at Week 16 regardless of the occurrence of intercurrent events. This analysis will use the same models specified for the primary analysis, where study participants are analyzed according to their randomized treatment, even if they discontinued study treatment for any reason prior to Week 16. Even though efforts will be made to collect the primary outcome data for all study participants at Week 16, there may still be some study participants for whom Week 16 efficacy data cannot be obtained. In this case, missing data will be imputed using multiple imputation under the assumption of MAR (see Section 4.2.2). The same analysis model as in the primary efficacy analyses will then be used on the imputed data set and the resulting inferential statistics will then be combined into a single inference using Rubin's rule.

8.2.3.5 Analysis on observed cases

An additional supportive analysis will be based on observed data only for study participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data at Week 16 will be treated as missing (see Section 4.2.2).

The same analysis model as in the primary efficacy analyses will then be used on the imputed data set.

8.2.3.6 Analysis on FAS

The primary efficacy analyses from Section 8.2.2 will be repeated based on the FAS.

8.2.3.7 Analysis on PPS

The primary efficacy analyses from Section 8.2.2 will be repeated based on the PPS.

8.2.3.8 Analysis on CFS

The primary efficacy analyses from Section 8.2.2 will be repeated based on the CFS.

8.2.3.9 Analysis including COVID-19 impact as intercurrent event

An additional sensitivity analysis will include an additional intercurrent event. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and completion of study treatment through Week 16 without receiving systemic antibiotic rescue medication or discontinuing IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for this sensitivity efficacy analysis:

1. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
2. Study participant-level outcome=HiSCR₅₀ at Week 16.
3. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication, discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16, or missing data due to COVID-19. More information is provided in Section 3.9. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, not discontinuing study treatment due to an AE or lack of efficacy through Week 16, and not having missing data due to COVID-19. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation as defined in Section 4.2.1.
4. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

The same imputation techniques and analysis model as in the primary efficacy analyses will then be used.

8.2.3.10 Cochran-Mantel-Haenszel test

The primary efficacy analyses from Section 8.2.2 will be repeated where the CMH test with fixed effects for treatment, Hurley stage at Baseline, and Baseline antibiotic use will be used as stratification variables. Pairwise treatment comparisons will be made based on the CMH test using the p-value for the general association. The odds ratio and associated confidence interval (CI) based on the Wald test will be provided.

8.2.3.11 Center-by-Treatment Interaction

The center-by-treatment interaction will be tested by adding center and a center-by-treatment interaction term in the logistic regression model described in Section 8.2.2. In the model, center will be based on the original centers prior to pooling. However, if the model is unable to converge due to a low number of participants at a given center, a pooling (see Section 3.7) will be described in order to allow the model to converge. In order to obtain reasonable estimates of variability for a treatment arm at a given center, a minimum of 21 participants will be considered acceptable for a center to be included in the model without pooling. Given the 2:2:2:1 randomization allocation scheme, this should provide a minimum of about 12 participants in the bimekizumab 320mg Q2W treatment group, 6 participants in the bimekizumab 320mg Q4W treatment group, and 3 participants in the placebo treatment group. Centers with fewer than 21 participants will be eligible for pooling. The pooling algorithm used is described in Section 3.7.

In order to achieve model convergence, other explanatory variables eg, Baseline Hurley Stage and Baseline antibiotic use may be dropped from the model. If model convergence is still not achieved, region and a region-by-treatment interaction term will be added to the model instead. Regions are defined in Section 3.7.

If the center-by-treatment interaction is not found to be significant ($\alpha=0.05$), then no further analyses will be performed. On the other hand, if the interaction is significant, further analyses

will be conducted to determine which center or centers may be the source of interaction. This will be done by running the logistic regression model (including the interaction term) where each center will be systematically removed from the model. This impact of a given center will be based on the change in the interaction p-value when that center is removed. The center or centers that appear to be driving the significant interaction effect will then be removed from the model to verify that conclusions remain the same with or without the influential center(s). This sensitivity analysis will be based on RS with MI/MCMC Monotone Regression for missing data.

8.3 Secondary efficacy endpoints

The secondary efficacy analyses will be performed based on the RS. Sensitivity analyses of the secondary endpoints will be performed on the CFS.

Missing data handling and sensitivity analyses of the secondary efficacy endpoints are described in Section 4.2.1.2.

The analyses of the secondary endpoints are summarized in Table 8–2 .

Table 8–2: Estimand Details and Attributes for Secondary Analyses

		Estimands for Secondary Endpoints			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Secondary Objective: Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS					
HiSCR ₇₅	Secondary (Section 8.3.1)	HiSCR ₇₅ response at Week 16	RS	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
Flare	Secondary (Section 8.3.2)	Flare by Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (flare).	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.
DLQI	Secondary (Section 8.3.3.1)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
DLQI	Secondary - Sensitivity (Section 8.3.3.2)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a DLQI total score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.
HSSDD	Secondary (Section 8.3.4.1)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.4.2)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.
HSSDD	Secondary (Section 8.3.5.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.

Table 8–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary Sensitivity (Section 8.3.5.2.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.
HSSDD	Secondary Sensitivity (Section 8.3.5.2.2)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

AE=adverse event; ANCOVA=analysis of covariance; DLQI=Dermatology Life Quality Index; HiSCR=Hidradenitis Suppurativa Clinical Response; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; IES=intercurrent event(s) strategy; LSMD=Least Squares Mean Difference; MCMC=Markov Chain Monte Carlo; MI=multiple imputation; PLS=Population-level summary; Pop=Population; RS=Randomized Set

^a Analysis includes all study participants in the RS with a Baseline HSSDD Worst Skin Pain score of 3 or higher.

8.3.1 HiSCR₇₅ at Week 16

A categorical response variable, HiSCR₇₅ at Week 16 is defined to be equal to 1 if %ΔAN is less than or equal to -75%, and the change from Baseline in the abscess count is zero or negative (ie, non-positive) together with a non-positive change from Baseline in the draining tunnel (fistula/sinus tract) count, and 0 otherwise. This definition is introduced for identifying participants who respond to the treatment (1 = responder, 0 = nonresponder). The definition of percentage improvement from Baseline is given in Section 8.2.1.

For HiSCR₇₅ at Week 16, logistic regression as specified for the primary analysis will be implemented to test for superiority. The same analysis approach as outlined for the primary efficacy endpoint will be applied.

8.3.2 Flare by Week 16

See Section 8.2.1 for the derivation of AN count.

Disease flare by Week 16 is defined as at least a 25% increase in AN count with an absolute increase of ≥ 2 AN relative to Baseline is observed by Week 16. A participant's disease flare status (yes/no) will be determined at each visit.

The number of participants who experience at least 1 disease flare by Week 16 will be analyzed using a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, and Baseline antibiotic use. The odds ratio versus placebo, p-value (from Wald test), and CI will be calculated. Missing data will be handled as described in Section 4.2.1.2.

8.3.3 Change from Baseline in DLQI Total Score at Week 16

The DLQI is a questionnaire designed for use in adult participants with skin diseases and has been used in patients with HS. This is a validated, quality-of-life questionnaire that covers 6 domains including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment, as assessed over the past week.

The scoring of each answer for the DLQI is as follows:

Table 8–3: Dermatology Life Quality Index

DLQI Scoring	
Response	Score
Very much	3
A lot	2
A little	1
Not at all	0
Not relevant	0
Question unanswered	0
Q7: ‘prevented work or studying’ = yes	3

The DLQI total score is calculated by adding the score of each question. The maximum score is 30, and the minimum score is 0. The higher the score, the more quality of life is impaired.

Meaning of DLQI Total Score

0-1 = no effect at all on patient’s life

2-5 = small effect on patient’s life

6-10 = moderate effect on patient’s life

11-20 = very large effect on patient’s life

21-30 = extremely large effect on patient’s life

This categorization will not be utilized in the analysis.

Because Q7 has a sub-question (referred to as Q7a here) after the leading yes/no question, some clarifying rules for scoring are provided:

- If Q7 is marked as “yes”, a score of 3 is given regardless of the responses to Q7a.
- If Q7 is marked as “no”, “not relevant”, or is missing and Q7a is “A lot”, a score of 2 is given.
- If Q7 is marked as “no”, “not relevant”, or is missing and Q7a is “A little”, a score of 1 is given.
- If Q7 is marked as “no”, “not relevant”, or is missing and Q7a is “Not at all”, a score of 0 is given.
- If Q7 is marked as “no” or “not relevant” and Q7a is missing, a score of 0 is given.
- If Q7 is missing and Q7a is missing, Q7 is considered unanswered (see below for details on how this impacts the DLQI total score).

If 1 question is left unanswered, this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30. If 2 or more questions are left unanswered, the questionnaire is not scored.

Change from Baseline in DLQI total score is defined as Week 16 DLQI total score minus Baseline DLQI total score.

8.3.3.1 Primary analysis of change from Baseline in DLQI Total score at Week 16

Missing data imputation described in Section 4.2.1.2 will be applied.

Change from Baseline in DLQI total score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use and Baseline value as a covariate. The least square mean (LSM), standard error (SE), and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab: the difference between the LSM, the associated 97.5% CI for the contrasts, and the corresponding p-value will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96. Estimand and intercurrent event details are specified in Table 8–2.

8.3.3.2 Sensitivity analysis of change from Baseline in DLQI Total score at Week 16

A sensitivity analysis using the same analysis model as in Section 8.3.3.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

8.3.4 Change from Baseline in Skin Pain score at Week 16, as assessed by the “worst skin pain” item in the HSSDD

The items on the HSSDD assess patients’ perception of the core symptoms of HS experienced in the past 24 hours: pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point numeric rating scale (NRS). Two items assess skin pain: worst skin pain and average skin pain.

Weekly averages will be derived for each of the items of the HSSDD for weeks matching the post-Baseline dosing weeks up to Week 16. A weekly average is the sum of the scored item over the course of the study week divided by the number of days on which the item was completed. The weekly averages will be relative to the respective visit date except for Baseline, which will be anchored to the first dose of study drug. A weekly average will only be calculated if at least 4 non-missing values (not necessarily consecutive) are available. Otherwise, the HSSDD weekly average for the given question will be set to missing.

Baseline will be computed as the average from the first 7 consecutive day period in which there are at least 4 non-missing entries. That is, first consider the first 7 consecutive days prior to the Baseline visit, but not including the Baseline visit day itself. If there are at least 4 non-missing values (not necessarily consecutive), then the Baseline average will be calculated. If there are less than 4 values, the 7 consecutive day period will move one day earlier. If there are at least 4 non-missing values (not necessarily consecutive) in that period, then the Baseline average will be calculated. This will continue until there are at least 4 non-missing values in a 7 consecutive day

period in the 14 days prior to Baseline. If there is no period in which there are at least 4 non-missing entries, then the Baseline value will be set to missing.

Change from Baseline in worst skin pain score is defined as the average Week 16 worst skin pain score minus the Baseline worst skin pain score. Missing data imputation described in Section 4.2.1.2 will be applied to the weekly averages and not to the individual daily PRO data.

8.3.4.1 Primary analysis of change from Baseline in skin pain score at Week 16

Change from Baseline in worst skin pain score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use, analgesic use (Section 6.4.2) and Baseline value as a covariate.

The LSM, SE, and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab, the difference between the LSM, the associated 97.5% CI for the contrasts, and the corresponding p-value will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.

8.3.4.2 Sensitivity analysis of change from Baseline in skin pain score at Week 16

A sensitivity analysis using the same analysis model as in Section 8.3.4.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

8.3.5 HSSDD skin pain response at Week 16

The analysis set for the analyses of the skin pain response will be restricted to those study participants in the RS with a Baseline worst skin pain score of 3 or higher. The weekly scores and Baseline score are derived as specified in Section 8.3.4.

8.3.5.1 Primary analysis of skin pain response at Week 16

Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, is defined as an improvement in the weekly worst skin pain score of at least 3 points versus Baseline.

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, Baseline antibiotic use, and analgesic use (Section 6.4.2).

The odds ratio versus placebo, p-value (from Wald test), and 97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose. Missing data will be handled as specified in Section 4.2.1.2. Estimand and intercurrent event details are specified in Table 8–2.

The number and percentage of participants who are pain responders at Week 16 will be summarized by treatment group.

By-participant listings of pain response status will be provided.

8.3.5.2 Sensitivity analyses of Skin Pain Response at Week 16

8.3.5.2.1 Nonresponse imputation

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (Table 8–2) will be imputed as nonresponse. That is, participants who experience an intercurrent event will be imputed as nonresponder at the timepoint of the event and all subsequent timepoints (including any recorded data after the event), and all missing data will also be imputed as nonresponse.

The same analysis model as Section 8.3.5.1 will then be used on the imputed data set.

8.3.5.2.2 Analysis on observed cases

An additional supportive analysis will be based on observed data only for study participants **with a worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16**. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing (see Section 4.2.2).

The same analysis model as in Section 8.3.5.1 will then be used on the imputed data set.

8.4 Other efficacy endpoints

The other efficacy endpoints are listed below and will be evaluated according to the planned assessments in the protocol. This excludes the timepoints for the primary and secondary endpoints specified above in Section 8.2.1 and Section 8.3.

Missing data handling for these endpoints is described in Section 4.2.1.3.

8.4.1 HiSCR endpoints

8.4.1.1 HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀

Categorical response variables HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ are defined to be equal to 1 if %ΔAN is less than or equal to -25%, -50%, -75%, -90%, and -100%, respectively, and the change from Baseline in the abscess count is zero or negative (ie, non-positive) together with a non-positive change from Baseline in the draining tunnel (fistula/sinus tract) count, and 0 otherwise. This definition is introduced for identifying study participants who respond to the treatment (1 = responder, 0 = nonresponder). The definition of percentage improvement from Baseline is given in Section 8.2.1.

HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response will be summarized using frequency tables by treatment group and visit.

A line plot of the HiSCR responder (HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀) rate over time, by treatment group, will be produced.

In order to investigate the effect of intercurrent event handling and missing data handling on the binary response variables, the following iterations of each of the aforementioned plots (HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀, respectively) will be presented, with corresponding summary statistics tables:

- Observed data only (non-imputation for either intercurrent events or missing data)

- Non-response imputation to reflect intercurrent events, non-imputation for missing data
- Full imputation: non-response imputation to reflect intercurrent events, imputation for missing data per Section 4.2.1.3

Bar charts of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ will be produced by visit and Hurley Stage at Baseline for the Initial Treatment Period and the combined Initial and Maintenance Treatment Period. These bar charts will be repeated for HiSCR by visit and Baseline antibiotic use.

Another bar chart of HiSCR rate at Week 16 will be produced by Hurley Stage at Baseline and Baseline antibiotic use.

In addition to the above bar charts, a stacked bar chart displaying whether the criteria are met (yes/no) for each of the component data used in the calculation of HiSCR (%ΔAN, abscess count and draining tunnel [fistula/sinus tract] count) will be generated by treatment group and visit. This graph will summarize the proportion of participants in each of the 8 different yes/no binary responses at each visit (2 x 2 x 2 response combination) at each visit.

8.4.1.2 **Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀**

See Section 8.4.1.1 for the derivation of HiSCR.

Initial Treatment Period

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first HiSCR_{xx} response, Date of Week 16 visit) – Date of first dose of study medication + 1, here xx represents 25, 50, 75, 90, 100 respectively. All visits including unscheduled visits are considered.

Participants who discontinue study treatment without achieving a given HiSCR response prior to Week 16 visit will be censored at the date of last lesion count assessment. Participants who reach the Week 16 Visit without achieving the given response will be censored at the date of the Week 16 Visit. Participants who experience an intercurrent event prior to achieving a HiSCR response will be censored at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during Initial Treatment Period will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to HiSCR responses will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

Combined Initial and Maintenance Treatment Period

An additional time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the combined Initial and Maintenance Treatment Period will be calculated as above, where the Week 48 visit is considered instead of Week 16.

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the combined Initial and Maintenance Treatment Period will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to HiSCR responses will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.

8.4.1.3 HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response at both Weeks 16 and 48

See Section 8.4.1.1 for the derivation of HiSCR response.

The number and percentage of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at both Weeks 16 and 48 will be summarized based on the RS and MS.

Missing data for the above summaries will be handled using NRI. That is, participants are counted as responders only if they have an observed HiSCR at both Weeks 16 and 48 and have no intercurrent events through Week 48. Otherwise, they are treated as not responding.

8.4.1.4 HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ among Week 16 Responders

See Section 8.4.1.1 for the derivation of HiSCR response.

Summaries of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at each visit from Week 16 through Week 48 will be summarized based on a subset of participants in the MS who achieve response at Week 16. The summaries will be as follows:

- HiSCR₅₀ responder rate based on participants who achieved HiSCR₅₀ response at Week 16
- HiSCR₇₅ responder rate based on participants who achieved HiSCR₇₅ response at Week 16
- HiSCR₉₀ responder rate based on participants who achieved HiSCR₉₀ response at Week 16
- HiSCR₁₀₀ responder rate based on participants who achieved HiSCR₁₀₀ response at Week 16

Line plots of the above HiSCR responder rate categories over time (from Week 16 to Week 48), by treatment group, will be produced.

8.4.1.5 Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders

See Section 8.4.1.1 for the derivation of HiSCR response.

Time to loss of response will be based on the MS and include only participants who had the corresponding HiSCR response at Week 16 (considering intercurrent event handling from the composite estimand described in Section 8.2.2).

Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) is defined as: Date of loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ - Date of Week 16 treatment administration + 1.

Time to loss of response will be estimated and presented using the Kaplan-Meier product-limit method for each treatment. Participants who experience an intercurrent event prior to loss of response will be considered as having lost response on the date of intercurrent event. Participants who reach the Week 48 Visit without loss of response will be censored at the date of the Week 48 Visit. Participants who discontinue treatment or study, for reasons other than those already defined for an intercurrent event, and who have not yet displayed loss of response by the time of withdrawal, will be censored at the date of the last lesion count assessment.

8.4.1.6 Partial response

See Section 8.2.1 for the derivation of AN count.

A partial response is defined as a $\geq 25\%$ reduction in AN count from Baseline (Week 0) at a particular timepoint.

The number and percentage of participants who are partial responders at Week 16 and become HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders in the Maintenance Treatment Period will be summarized by treatment group and visit. These analyses will be based on the subset of participants in the MS that are partial responders but not HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders, respectively, at Week 16. These summaries will be based on observed case data and will not consider the occurrence of intercurrent events.

8.4.2 Lesion count

8.4.2.1 Change from Baseline in lesion count

At each visit, lesion counts will be summarized by treatment group, anatomical region, and lesion type. The following lesion types will be summarized:

- Abscesses
- Inflammatory nodules
- ANs
- Non-inflammatory nodules
- Draining tunnels (fistulas/sinus tracts)
- Non-draining tunnels (fistulas/sinus tracts)
- HS scars

Total lesion counts (ie, the total across all anatomical regions) will be summarized by visit and lesion type, treatment group, overall and by Baseline Hurley Stage. Summaries will also be presented for the change and percentage change from Baseline in lesion counts by anatomical region and total lesion counts by lesion type.

A line plot of the percentage change from Baseline in AN count over time by treatment group will be produced. Separate plots will also be produced for the percentage change from Baseline

in the abscess count, inflammatory nodules count, and draining tunnel count, respectively, over time.

In order to investigate the effect of intercurrent events and missing data handling on lesion count data, the following **additional** iterations of each of the aforementioned plots (percentage change in AN count, abscess count, inflammatory nodule count and draining tunnel count, respectively) will be presented, with corresponding summary statistics tables:

- Intercurrent events:
 - Participants who have experienced intercurrent events
 - Participants who have not experienced intercurrent events
- Missing data handling:
 - Observed lesion counts (i.e., non-imputation)
 - Imputed lesion counts

Lesion count data will be listed by treatment group and anatomical region and will show region-specific Hurley Stage and worst overall Hurley Stage for each participant and visit. The total count for each type of lesion, across all anatomical regions at each visit will be listed separately. For the total abscess count and total draining tunnel (fistula/sinus tract) count, the change from Baseline will be listed; for the AN count, the percentage change from Baseline will be listed.

8.4.2.2 AN count of 0, 1, or 2

The number and percentage of participants with an AN count of 0, 1, or 2 will be presented by treatment group and visit. The denominator for the percentage calculations will be the number of participants in each treatment group in the RS with non-missing data at each visit.

8.4.2.3 AN₂₅, AN₅₀, AN₇₅, AN₉₀, AN₁₀₀

Categorical response variables AN₂₅, AN₅₀, AN₇₅, AN₉₀, and AN₁₀₀ are defined to be equal to 1 if %ΔAN is less than or equal to -25%, -50%, -75%, -90%, and -100%, respectively, and 0 otherwise. This definition is introduced for identifying participants who respond to the treatment (1 = responder, 0 = nonresponder). The definition of percentage change from Baseline is given in Section 8.2.1.

AN₂₅, AN₅₀, AN₇₅, AN₉₀, and AN₁₀₀ response will be summarized using frequency tables by treatment group for each visit.

8.4.3 Flare relative to Baseline

See Section 8.3.2 for the derivation of flare.

Disease flare status will be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in each treatment group. This summary will also include the number of participants with any flare in the Initial Period, Maintenance Period, and the combined Initial and Maintenance Period. A bar chart of percentage of participants with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Initial Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Initial Treatment Period will be presented.

8.4.4 Time to flare by Week 16

See Section 8.3.2 for the derivation of flare.

Time to flare (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first flare, Date of Week 16 visit) – Date of first dose of study medication + 1.
All visits in the Initial Treatment Period including unscheduled visits are considered.

Participants who discontinue study treatment without experiencing a flare prior to Week 16 Visit will be censored at the date of last lesion count assessment. Participants who reach the Week 16 Visit without experiencing a flare will be censored at the date of the Week 16 Visit. Participants who experience an intercurrent event prior to experiencing a flare will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.

The median time to flare, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

8.4.5 Time to flare by Week 48

See Section 8.3.2 for the derivation of flare relative to Baseline.

Time to flare (in days) during the combined Initial and Maintenance Treatment Period will be calculated as:

Min (Date of first flare, Date of Week 48 visit) – Date of first dose of study medication + 1.
All visits in the up to Week 48 including unscheduled visits are considered.

Flare will be defined relative to the Baseline visit. Participants who discontinue study treatment without experiencing a flare prior to Week 48 visit will be censored at the date of last lesion count assessment. Participants who reach the Week 48 Visit without experiencing a flare will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event prior to experiencing a flare will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.

The median time to flare, including the 2-sided 95% confidence interval, will be calculated for each treatment.

8.4.6 International Hidradenitis Suppurativa Severity Score System (IHS4)

The IHS4 is a validated scoring tool to dynamically assess HS severity to be used both in real-life and clinical trials settings (Zouboulis et al, 2017). The determination of the IHS4 score requires counting the inflammatory nodules, abscesses and draining tunnels (fistulas/sinus tracts), making it straightforward to apply in both research and clinical practice and easy to use in conjunction with the HiSCR.

$$IHS4 = (number\ of\ inflammatory\ nodules \times 1) + (number\ of\ abscesses \times 2) + (number\ of\ draining\ tunnels\ (fistulas/sinus\ tracts) \times 4)$$

The IHS4 score will be derived based on observed component total lesion count data; in the case of missing component data, the IHS4 score will be missing.

The observed IHS4 score, change and percentage change from Baseline will be summarized by treatment group and visit. Missing IHS4 scores will be imputed using the multiple imputation procedure specified in Section 4.2.2.1, where IHS4 scores will be derived based on the imputed lesion counts.

The IHS4 scores will be categorized into 3 HS categories (mild HS: ≤ 3 , moderate HS: 4-10, severe HS: ≥ 11).

The number and percentage of participants in each category (mild, moderate, severe) will be presented by treatment group and visit. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there are no missing data (OC).

Shift tables for the changes from Baseline in this scale will be presented for each post-Baseline visit by treatment group. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there is no missing data for the change (OC).

8.4.7 HS-Physician's Global Assessment 6-point scale

The HS-Physician's Global Assessment is a validated 6-point scale that is used to measure improvement in inflammatory nodules, abscesses, and draining tunnels. The HS-Physician's Global Assessment scale is defined in Table 8-4.

Table 8-4: HS-Physician's Global Assessment 6-point scale

Score	Rating	Description
0	Clear	No abscesses, no draining tunnels (fistulas/sinus tracts), no nodules
1	Minimal	No abscesses, no draining tunnels (fistulas/sinus tracts), no inflammatory nodules, presence of non-inflammatory nodules

Table 8–4: HS-Physician's Global Assessment 6-point scale

Score	Rating	Description
2	Mild	No abscesses or draining tunnels (fistulas/sinus tracts), and less than 5 inflammatory nodules, or Single abscess or draining tunnel (fistula/sinus tract), and no inflammatory nodules
3	Moderate	No abscesses or draining tunnels (fistulas/sinus tracts), and at least 5 inflammatory nodules, or Single abscess or draining tunnel (fistula/sinus tract) in the presence of inflammatory nodules, or Between 2 and 5 abscesses or draining tunnels (fistulas/sinus tracts) with or without inflammatory nodules, up to 10
4	Severe	Between 2 and 5 abscesses and draining tunnels (fistulas/sinus tracts), with inflammatory nodules that are greater than 10
5	Very severe	More than 5 abscesses or draining tunnels (fistulas/sinus tracts)

The gradings will be derived on a participant-level basis (ie, across all anatomical regions) based on the following rules:

- Clear:
 - number of abscesses = 0;
 - number of draining tunnels (fistulas/sinus tracts) = 0;
 - number of inflammatory nodules = 0;
 - number of non-inflammatory nodules = 0;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit;
- Minimal:
 - number of abscesses = 0;
 - number of draining tunnels (fistulas/sinus tracts) = 0;
 - number of inflammatory nodules = 0;
 - number of non-inflammatory nodules ≥ 1 ;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit;
- Mild:
 - number of abscesses = 0;
 - number of draining tunnels (fistulas/sinus tracts) = 0;

- number of inflammatory nodules ≥ 1 and ≤ 4 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

OR

- sum of number of abscesses and number of draining tunnels (fistulas/sinus tracts) = 1;
- number of inflammatory nodules = 0;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

- Moderate:

- number of abscesses = 0;
- number of draining tunnels (fistulas/sinus tracts) = 0;
- number of inflammatory nodules ≥ 5 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

OR

- sum number of abscesses and number of draining tunnels (fistulas/sinus tracts) = 1;
- number of inflammatory nodules ≥ 1 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

OR

- sum of number of abscesses and number of draining tunnels (fistulas/sinus tracts) ≥ 2 and ≤ 5 ;
- number of inflammatory nodules ≤ 10 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;

- Severe:

- sum of number of abscesses and number of draining tunnels (fistulas/sinus tracts) ≥ 2 and ≤ 5 ;
- number of inflammatory nodules > 10 ;
- number of non-inflammatory nodules – no limit;
- number of non-draining tunnels (fistulas/sinus tracts) – no limit;
- number of HS scars – no limit;
- Very Severe:
 - sum of number of and number of draining tunnels (fistulas/sinus tracts) > 5 ;
 - number of inflammatory nodule – no limit;
 - number of non-inflammatory nodules – no limit;
 - number of non-draining tunnels (fistulas/sinus tracts) – no limit;
 - number of HS scars – no limit

The number and percentage of participants at each level of the assessment scale (Clear, Minimal, Mild, Moderate, Severe and Very severe) will be presented by treatment group and visit. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there are no missing data for the HS-Physician's Global Assessment.

Shift tables for the changes from Baseline in this scale will be presented for each post-Baseline visit by treatment group. The denominator for the percentages will be based on the number of participants in the given treatment group and visit for which there is no missing data for the change from Baseline in HS-Physician's Global Assessment.

The HS-Physician's Global Assessment will be listed by treatment group and participant at each visit.

8.4.8 High Sensitivity C-Reactive Protein (hs-CRP)

Concentrations of hs-CRP, changes from Baseline, and percent change from Baseline will be summarized by treatment group and visit, where percent change is calculated as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline hs-CRP} - \text{Baseline hs-CRP}}{\text{Baseline hs-CRP}}$$

Summary statistics will include n, arithmetic mean, SD, median, Q1, Q3, minimum and maximum. For the ratio to Baseline, summary statistics will include n, geometric mean, geoCV, median, Q1, Q3, minimum and maximum.

The ratio to Baseline will be calculated as follows:

$$\text{Ratio to Baseline} = \text{hs-CRP at post-Baseline} / \text{hs-CRP at Baseline visit}$$

The ratio to Baseline will also be summarized by treatment group and visit.

For the hs-CRP data, measurements that are below the limit of quantification (BLQ) will be imputed with half of the lower limit of quantification (LLOQ) for the purpose of calculating summary statistics, changes from Baseline, and ratio to Baseline.

Concentrations of hs-CRP, changes from Baseline, and ratio to Baseline will be listed.

8.4.9 Initiation of systemic antibiotic rescue therapy

See Section 3.9 for the definition of a systemic antibiotic rescue therapy.

The number of participants that use rescue antibiotic therapy will be summarized by treatment group for each period.

8.4.10 Time to initiation of systemic rescue therapy in the Initial Treatment Period

See Section 3.9 for the definition of a systemic antibiotic rescue therapy.

Time to initiation of systemic rescue therapy (in days) during the Initial Treatment Period will be calculated as:

Min (Date of initiation of rescue therapy, Date of change in the dose/type of current antibiotic, Date of Week 16 visit) – Date of first dose of study medication + 1.

Participants who discontinue the study without initiating systemic rescue therapy prior to Week 16 visit will be censored at the date of discontinuation. Participants who reach the Week 16 Visit without initiating systemic rescue therapy will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no Post-Baseline visit.

Time to initiation of systemic rescue therapy will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to initiation of systemic rescue therapy will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to requiring rescue therapy.

The median time to initiation of systemic rescue therapy, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

8.4.11 Time to an intercurrent event in the Initial Treatment Period

See Section 3.9 for the definition of an intercurrent event.

Time to an intercurrent event (in days) during the Initial Treatment Period will be calculated as:

Min (Date of intercurrent event, Date of Week 16 visit) – Date of first dose of study medication + 1.

Participants who discontinue the study without experiencing an intercurrent event prior to Week 16 visit will be censored at the date of discontinuation. That includes participants who discontinue from the study for reasons other than Adverse Event and Lack of Efficacy. Participants who reach the Week 16 Visit without experiencing an intercurrent event will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no Post-Baseline visit.

Time to an intercurrent event will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to initiation of systemic rescue therapy will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to intercurrent event.

The median time to an intercurrent event, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

8.4.12 Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)

See Section 8.3.4 for details on HSSDD Baseline and weekly average definitions and derivations.

Percent change from Baseline in HSSDD responses for worst and average skin pain score is defined as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline HSSDD score} - \text{Baseline HSSDD score}}{\text{Baseline HSSDD score}}$$

Change from Baseline in each HSSDD item (worst skin pain, average skin pain, smell or odor, itch at its worst, and amount of drainage or oozing) score will be summarized using descriptive statistics by treatment group and visit, based on weekly averages. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits. Percentage change will be summarized for the worst and average skin pain items.

Additionally, change from Baseline in each HSSDD item will be evaluated by treatment group at Week 16 via continuous empirical cumulative distribution function (eCDF) plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Change from Baseline in Worst Skin Pain score and Worst Itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

HSSDD response based on clinically meaningful change for the worst skin pain item is defined as at least a 3-point reduction from Baseline in HSSDD among study participants with a score of ≥ 3 at Baseline, based on weekly averages.

The number and percentage of responders based on clinically meaningful change for the worst skin pain item will be summarized by treatment group and visit.

The number and percentage of participants who were responders based on clinically meaningful change at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst skin pain item.

HSSDD response for the worst skin pain and average skin pain items is defined as at least a 30% reduction and at least a 1-point reduction from Baseline among study participants with a score of ≥ 3 at Baseline. The number and percentage of responders for each item will be summarized by treatment group and visit.

The number and percentage of participants who were responders (based on the 30% improvement and 1 point improvement definition) at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst skin pain and average skin pain items.

The number and percentage of participants that complete the HSSDD will be calculated for each visit by treatment group. A participant will be counted as completing the HSSDD at a visit if the minimum number of daily entries is present to calculate the weekly average (see Section 8.3.4). The percentage will be based on the number of participants in the RS. A participant will be considered a completer at a visit if the weekly average can be calculated for that visit.

8.4.13 Hidradenitis Suppurativa Symptom Questionnaire (HSSQ)

The 4 items on the HS Symptom Questionnaire (HSSQ) assesses participants' perception of the core symptoms of HS experienced in the past 7 days - skin pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point NRS.

The change from Baseline score is derived as post Baseline score minus Baseline score. A negative change score indicates a reduction in the score/improvement for the participant.

Summary statistics of the actual values and change and percentage change from Baseline values will be used to summarize each HSSQ item for each visit by treatment group. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Additionally, change from Baseline in each HSSQ item will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

HSSQ response for skin pain item is defined as at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline.

The number and percentage of responders for skin pain item will be summarized by treatment group and visit based on the MS.

The number and percentage of participants who were responders at any timepoint in the Maintenance Treatment Period will be summarized by treatment group for the skin pain score based on the MS.

Change from Baseline in skin pain score and itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

The number and percentage of participants that complete the HSSQ will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the MS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if each of the items are completed at that visit.

8.4.14 DLQI

See Section 8.3.3 for the derivation of DLQI total score.

A DLQI total score of 0 or 1 indicates no impact of the skin disease on health-related quality of life and will be summarized.

A participant is considered to have achieved the minimally clinical important difference (MCID) if their individual improvement (ie, decrease) from Baseline in total score is ≥ 4 . A 4-point improvement in the DLQI total score (DLQI response) has been reported to be meaningful for the participant (within-participant MCID). The summary of MCID will be restricted to participants with a DLQI total score of at least 4 at Baseline to ensure that it is possible for the participant to achieve the MCID.

The DLQI related efficacy variables are defined as follows:

- Change from Baseline in DLQI total score is defined as Post-Baseline DLQI total score minus Baseline DLQI total score.
- Percent of study participants achieving a DLQI total score of 0 or 1 is defined as the number of study participants with DLQI total score of 0 or 1 divided by the number of study participants in RS.
- Percent of study participants achieving a MCID in DLQI total score is defined as the number of study participants with improvement from Baseline in total score of 4 or more divided by the number of study participants in RS that have a Baseline DLQI total score of at least 4.

Missing data for the DLQI total score will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

Change from Baseline in DLQI total score will be summarized using descriptive statistics by treatment group and visit. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Frequency tables will be produced to show the number and percentage of DLQI responders for MCID for each visit by treatment groups.

The number and percentage of participants achieving a DLQI total score of 0 or 1 at each visit will be summarized descriptively using counts and percentages by treatment group and visit.

The number and percentage of participants that complete the DLQI total score will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the total score is calculated at that visit.

A by-participant listing of the DLQI questionnaire, DLQI total score, change from Baseline and DLQI response for MCID and 0 or 1 data will be provided by treatment group.

8.4.15 Hidradenitis Suppurativa Quality of Life (HiSQOL)

The HiSQOL includes 17 items assessed using a 7-day recall period, grouped in 3 subscales: symptoms, psychosocial, activities and adaptation.

The assessment of each item of the HiSQOL is as follows:

Table 8–5: Hidradenitis Suppurative Quality of Life

HiSQOL Scoring	
Response	Score
Unable to do, due to my HS	4
Extremely	4
Very Much	3
Moderately	2
Slightly	1
Not at all	0
I normally do not do this, HS did not influence	0
I am not sexually active	0
I do not work or study	0
Unanswered	0

The HiSQOL total score is calculated by adding the score of each question. The maximum score is 68, and the minimum score is 0.

Subscale scores will be summarized for each of the 3 subscales. The maximum scores for the subscales are 16 (symptoms), 20 (psychosocial), and 32 (activities and adaptations), and the minimum score is 0 for all subscales.

For all scores, the higher the score, the more quality of life is impaired.

Summary statistics of the actual values and change from Baseline values will be used to summarize HiSQOL domain and total scores for each visit by treatment group. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Additionally, change from Baseline in each HiSQOL subscale will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3. The imputed HiSQOL total score will be derived based on the imputed subscales.

The number and percentage of participants that complete the HiSQOL will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the total score is calculated at that visit.

A by-participant listing of the HiSQOL questionnaire, HiSQOL responses, domain and total scores and change from Baseline will be provided.

8.4.16 Patient Global Impression of HS Severity (PGI-S-HS)

The PGI-S-HS is a single item to assess study participants' perceptions of the overall severity of HS over the past 7 days (none, mild, moderate, severe, very severe).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-S-HS will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.17 Patient Global Impression of Change in HS Severity (PGI-C-HS)

The PGI-C-HS is a single item to assess study participants' perception of the change in HS since they started taking the study medication (much better, a little better, no change, a little worse, much worse).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-C-HS will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.18 Patient Global Impression of Severity of Skin Pain (PGI-S-SP)

The PGI-S-SP is a single item to assess study participants' perceptions of the severity of their skin pain from their HS lesions, over the past 7 days (none, mild, moderate, severe, very severe).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-S-SP will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.19 Patient Global Impression of Change in Severity of Skin Pain (PGI-C-SP)

The PGI-C-SP is a single item to assess study participants' perceptions of change in their skin pain from their HS lesions, since they started taking the study medication (much better, a little better, no change, a little worse, much worse).

The number and percentage of participants with each response will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the PGI-C-SP will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the response is nonmissing at that visit.

8.4.20 Euro-Quality of Life 5-Dimensions, 3 levels (EQ-5D-3L)

The EQ-5D-3L comprises 5 questions on mobility, self-care, pain, usual activities, and psychological status with 3 possible answers for each item (1=no problem, 2=moderate problem, 3=severe problem). In addition, there is a VAS to indicate the general health status with 100 indicating the best health status.

Change from Baseline in EQ-5D-3L VAS scores is defined as Post-Baseline EQ-5D-3L VAS score minus Baseline EQ-5D-3L VAS score.

Responses to EQ-5D-3L will be summarized based on OC only as primary analysis. No imputation is applied to responses to EQ-5D-3L but is applied to EQ-5D-3L VAS scores.

Changes from Baseline in EQ-5D-3L VAS will be summarized using descriptive statistics by treatment group and visit. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

The number and percentage of participants with each response in the EQ-5D-3L will be summarized for each visit by treatment group based on OC data.

The number and percentage of participants that complete the EQ-5D-3L will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if each of the domains and VAS are completed at that visit.

8.4.21 Work Productivity and Activity Impairment Questionnaire–Specific Health Problem (WPAI-SHP) v2.0 adapted to HS scores

The WPAI-SHP V2.0 is a patient-reported questionnaire that assesses study participant's employment status, work absenteeism, work impairment while working (presenteeism), overall work, and daily activity impairment attributable to a specific health problem. It has been used in several clinical studies of biologic therapy in participants with plaque PSO.

Five out of 6 items of the WPAI-SHP are regrouped into the 4 dimensions, with scores expressed as percentage, where higher numbers indicate greater impairment and less productivity, ie, worse outcomes, as described in the WPAI-SHP scoring rules.

The scoring rules for the WPAI-SHP are as follows:

Questions:

- 1 = currently employed
- 2 = hours missed due to specified problem
- 3 = hours missed other reasons
- 4 = hours actually worked

- 5 = degree problem affected productivity while working
- 6 = degree problem affected regular activities

Scores:

- Percent work time missed due to problem: $[\text{Q2 hours}/(\text{Q2 hours} + \text{Q4 hours})] * 100$
- Percent impairment while working due to problem: $[\text{Q5 score}/10] * 100$
- Percent overall work impairment due to problem: $[\text{Q2 hours}/(\text{Q2 hours} + \text{Q4 hours}) + [(1 - (\text{Q2 hours}/(\text{Q2 hours} + \text{Q4 hours})) \times (\text{Q5 score}/10)] * 100$
- Percent activity impairment due to problem: $[\text{Q6 score}/10] * 100$

A negative number will indicate a reduction in the score/improvement for participants.

The change from Baseline score is derived as post Baseline score minus Baseline score. A negative change score indicates a reduction in the score/improvement for the participant.

Summary statistics of the actual values and change from Baseline values will be used to summarize WPAI-SHP for each visit by treatment group. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

The number and percentage of participants that complete the WPAI-SHP will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if the percentages in each dimension are calculated at that visit.

A by-participant listing of the WPAI-SHP questionnaire, WPAI-SHP domains and change from Baseline will be provided.

8.4.22 Treatment Satisfaction Questionnaire – Medication-9

The TSQM-9 is an abbreviated 9-item version of the TSQM, excluding the side effects of medication domain. The domains included in the TSQM-9 include effectiveness (3 items), convenience (3 items) and global satisfaction (3 items). The TSQM-9 domain scores range from 0 to 100 with higher scores representing higher satisfaction.

The scores for each measure are as follows:

- Global Satisfaction:
 - If no items are missing: $([\text{Sum}(\text{Item 7 to Item 9}) - 3]/14) * 100$
 - If either Item 7 or 8 is missing: $([\text{Sum}(\text{the two completed items}) - 2]/10) * 100$
 - If Item 9 is missing: $([\text{Sum}(\text{Item 7 and Item 8}) - 2]/8) * 100$
- Effectiveness
 - If no items are missing: $([\text{Item 1} + \text{Item 2} + \text{Item 3}) - 3]/18) * 100$

- If one item is missing: $([(\text{Sum}(\text{the two completed items}) - 2]/12) * 100$
- Convenience
 - If no items are missing: $([\text{Sum}(\text{Item 4 to Item 6}) - 3]/18) * 100$
 - If one item is missing: $([\text{Sum}(\text{the two completed items}) - 2]/12) * 100$

Frequency tables will be produced to summarize answers provided to each of the 9 items of the TSQM-9 at Weeks 16 and 48 by treatment group. Responses to TSQM-9 will be summarized based on OC. No imputation will be applied.

The number and percentage of participants that complete the TSQM-9 will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit. A participant will be considered a completer at a visit if each of the domains are completed at that visit.

A by-participant listing of TSQM-9 will be provided.

8.4.23 Lesion intervention

Investigators will have the option to perform interventions in the event an acutely painful lesion occurs that requires immediate intervention.

The following rules will be used to assign a lesion intervention to a study period:

- **Initial Treatment Period:** a lesion intervention will be assigned to the Initial Treatment Period if it has been performed between the first administration of IMP on Day 1 up to and including Week 16.
- **Maintenance Treatment Period:** a lesion intervention will be assigned to the Maintenance Treatment Period if it has been performed between Week 16 through the Week 48 visit.

Methods for dealing with partial dates are specified in Section 4.2.4.

A listing of participants who receive any lesion intervention will be provided.

The number and percentage of participants who receive at least 1 lesion intervention will be summarized by treatment group for the Initial Treatment Period, Maintenance Treatment Period, and the combined Initial and Maintenance Treatment Period.

The number and percentage of participants with 2, 3, and 4 or more lesion interventions will also be summarized by treatment group for the Initial Treatment Period, Maintenance Treatment Period, and the combined Initial and Maintenance Treatment Period.

The number and percentage of participants with 2 or more lesion interventions performed on the same lesion will be summarized by treatment group for the Initial Treatment Period, Maintenance Treatment Period, and the combined Initial and Maintenance Treatment Period.

8.5 Additional statistical analyses of other efficacy endpoints

For selected other efficacy variables, it is of interest to perform statistical tests and to calculate inferential statistics. As these tests are not part of the multiplicity-controlled procedure, the associated p-values are considered nominal and are not controlled for multiplicity.

For binary variables, the analysis will follow what was specified for the primary analysis of the primary endpoint and the corresponding p-value reported. Missing values will be imputed as for the primary analysis. For continuous variables, the MI – MCMC / Monotone Regression approach used for the secondary continuous endpoint will be applied for the imputation model. The analysis model will be as for the corresponding secondary continuous endpoint analysis, unless otherwise indicated.

Below is a list of variables for which these nominal p-values will be calculated. The results of these inferential tests will be presented in a single table summarizing the testing performed outside of the multiplicity-controlled testing procedure.

All tests will be for both 320mg Q2W vs Placebo and 320mg Q4W vs Placebo (tested separately) and will be performed for the Week 12 and Week 16 visits only. If the Week 16 visit test is already part of the controlled testing procedure in the primary or secondary analyses, only Week 12 is indicated here.

- HiSCR₅₀ at Week 12
- HiSCR₇₅ at Week 12
- HiSCR₉₀
- HiSCR₁₀₀
- Flare by Week 12
- Time to flare by Week 12 (based on time-to-event analysis per Section 8.4.4 and adjusted appropriately)
- IHS4 change from Baseline
- IHS4 percentage change from Baseline
- HS Physician's Global Assessment: rate of participants who are Clear or Mild
- DLQI total score change from Baseline at Week 12
- Worst Skin Pain per HSSDD change from Baseline at Week 12
- Skin Pain response per HSSDD at Week 12

9 PHARMACOKINETICS AND PHARMACODYNAMICS

9.1 Pharmacokinetics

Pharmacokinetic variables will be analyzed for all participants in the PK-PPS. Bimekizumab plasma concentrations will be summarized for each treatment at each scheduled visit.

PK summaries will be based on observed values. No imputation will be used. However, if plasma concentration measurements are below the level of quantification (BLQ), then for calculation of the derived statistics the result will be set to ½ of the lower level of quantification (LLOQ). Descriptive statistics including geometric mean, geometric coefficient of variation, and geometric mean 95% CI if applicable will be calculated if at least 2/3 of the values of interest are above the LLOQ. If this is not the case, only median, minimum, and maximum will be presented.

Geometric mean plasma concentration will be plotted by treatment group, and by cumulative antibody status for participants randomized to bimekizumab on linear and log linear scale. In addition, HiSCR₅₀ response (0=not achieved, 1=achieved) assessed at each PK visit (excluding Week 1) will be plotted against the visit's bimekizumab plasma concentration, separated by treatment group. Spaghetti plots of bimekizumab plasma concentrations by week from bimekizumab first dosing separated by treatment group and antibody status will be presented for participants with and without HiSCR₅₀ response at Week 16.

If the dosing for a visit is +/- 7 days out of window, then the plasma concentration from that visit and all subsequent visits will be excluded from the PK summary. In addition, if the PK sampling date is >1 day after the dosing date, then the plasma concentration from that visit will be excluded from the PK summary.

All PK concentrations collected will be listed irrespective of the dosing or sampling occurring out of window.

9.2 Pharmacodynamics

Not applicable.

9.3 Immunogenicity

9.3.1 Autoantibodies

Not applicable.

9.3.2 Anti-bimekizumab antibodies

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-tiered assay approach: screening assay, confirmatory assay, and titration assay. Samples confirmed as positive within the confirmatory assay will be further evaluated for the presence of neutralizing anti-bimekizumab antibodies specific to IL-17AA, IL-17FF or both. Samples will be taken at Baseline, then at study Weeks 4, 8, 12, 16, 20, 24, 36 and 48, and at PEOT and SFU timepoints.

ADAb samples are not analyzed when study participants are on a treatment other than bimekizumab. For study participants who switch from placebo to bimekizumab, samples are analyzed starting at the visit when the switch to bimekizumab occurs (Week 16). The sample at Week 16 will act as the Baseline for that treatment group.

The screening cut point will be used to determine the status of anti-bimekizumab antibodies in the test sample as Positive Screen (PS) or Negative Screen (NS). For samples presenting anti-bimekizumab antibody levels that are PS, a further confirmatory assay will be performed, and the result of which will be reported as either Positive Immunodepletion (PI) or Negative Immunodepletion (NI).

ADAb status for each sample will be derived as follows:

- Sample values that are either NS, or PS and NI and where the bimekizumab concentration is less than the validated ADAb assay drug tolerance limit will be defined as anti-bimekizumab antibody negative.
- Sample values that are either NS, or PS and NI and where the bimekizumab concentration exceeds the validated ADAb assay drug tolerance limit will be defined as inconclusive.

- Sample values that are PS and PI will be defined as ADAb positive (regardless of availability of a titer value)
- Missing or non-evaluable samples will be defined as missing

Positive immunodepletion samples will be titrated, and the ADAb titer (reciprocal dilution factor including minimum required dilution) will be reported. Subsequently, PI samples will also be subject to a neutralizing assay to evaluate the potential of ADAb to neutralize the target binding of bimekizumab (IL-17AA or IL-17FF or both) in vitro.

Cumulative ADAb status will be derived as follows:

The ADAb status (positive, negative or missing) will be considered in a cumulative manner at each time point.

A study participant will be counted positive from the first visit at which the study participant achieved a positive ADAb sample result to the end of the treatment period, regardless of any missing/inconclusive or negative ADAb sample result.

If a study participant has only negative ADAb samples or only one missing/inconclusive sample with all other ADAb samples being negative, the study participant will be classified as negative. An exception remains for the Baseline Visit where only one sample could be available. If the sample is missing/inconclusive, then the sample will be classified as being negative for the cumulative ADAb status.

Otherwise, the study participant will be classified in the missing ADAb category.

Overall ADAb status will be derived as follows:

A study participant will be classified as:

- Positive if the study participant has at least one positive sample up to the time point of interest (regardless of having missing/inconclusive data).
- Negative if the study participant has all the samples negative or only one missing/inconclusive sample with negative ADAb samples up to the timepoint of interest.
- Missing if the study participant has more than one missing ADAb result (or have more than one inconclusive sample) and all other available ADAb samples are negative up to the time point of interest.

ADAb categories will be derived as follows:

- **Pre ADAb negative – treatment-emergent ADAb negative (Category 1):** Includes study participants who are anti-bimekizumab antibody negative at Baseline and anti-bimekizumab antibody negative at all sampling points during the period of interest (one post-Baseline missing/inconclusive sample is allowed for subjects with pre- anti-bimekizumab antibody negative sample). This group also includes study participants who have a missing or inconclusive sample (either missing or inconclusive or insufficient volume) at Baseline (ie, pre-treatment) with all post-Baseline samples as ADAb negative.
- **Pre ADAb negative – treatment-emergent ADAb positive (Category 2):** Includes study participants who are ADAb negative at Baseline and ADAb positive at any sampling points post-Baseline during the period of interest. This group also includes study participants who

have a missing sample (either missing or insufficient volume) at Baseline (ie, pre-treatment) with 1 or more post-Baseline samples as ADAb positive.

- **Pre ADAb positive – treatment-emergent reduced ADAb (Category 3):** Includes study participants who are ADAb positive at Baseline, and ADAb negative at all sampling points post-Baseline during the period of interest.
- **Pre ADAb positive – treatment-emergent unaffected ADAb positive (Category 4):** Includes study participants who are ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with titer values of the same magnitude as Baseline (ie, less than a predefined fold difference from the Baseline titer).
 - For this analysis, this is set at an increase of less than the validated Minimum Significant Ratio (MSR) of 2.07-fold from Baseline.
- **Pre ADAb positive – treatment-emergent ADAb boosted positive (Category 5):** Includes study participants who ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with increased titer values compared to Baseline (equal to or greater than a predefined fold difference increase from Baseline titer which will be defined within the validation of the assay).
 - For this analysis, this is set at an increase equal to or greater than the validated MSR of 2.07-fold from Baseline.
 - Note: for any study participant who is ADAb positive at Baseline and ADAb positive at a post-Baseline time point during the period of interest, but for whom titers are not available to determine treatment unaffected or treatment boosted status, the study participant will be considered as treatment boosted, assuming no other samples are available.
- **ADAb Inconclusive (Category 6):** Includes study participants who have an ADAb positive Baseline (pre-treatment) sample and some post-Baseline samples during the period of interest are missing or inconclusive, while other post-Baseline samples are ADAb negative.
- **Total treatment-emergent ADAb positivity (Category 7 [Categories 2 and 5 combined]):** Includes study participants who are pre ADAb negative – treatment-emergent ADAb positive (Category 2) and pre ADAb positive – treatment boosted ADAb positive (Category 5).
- **Total prevalence of pre- ADAb positivity (Category 8 [Categories 3, 4, 5 and 6 combined]):** Study participants that are tested ADAb positive at Baseline.
- **Missing:** Includes study participants who are ADAb negative, missing, or inconclusive at Baseline with some post-Baseline samples as missing or inconclusive, and other samples as ADAb negative.

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, the following categories can also be used:

- **ADAb positive** – This is defined as study participants who are anti-bimekizumab antibody positive on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- **ADAb negative** – Study participants for who either:

- All samples (including Baseline) are ADAb negative and there are no missing or inconclusive samples
- Only 1 sample is ADAb positive and all other samples (including Baseline) are ADAb negative or missing or inconclusive
- Only 1 sample is missing or inconclusive and the remaining ADAb samples are negative.
- ADAb missing - Defined as study participants who do not fulfil the criteria for one of the 2 groups listed above.

The rationale for requiring at least 2 time points in which ADAb levels are above the specified cut point is to exclude those study participants who have only one occurrence of ADAb levels during the course of treatment. Including such study participants would increase the number of ADAb positive study participants with potentially no impact on efficacy.

In the case that a sample is collected 1 or more days following the scheduled visit date in which the drug was administered, the ADAb results for that sample will be associated with the scheduled visit and summarized accordingly. Such samples will also be considered when ADAb results are summarized over a given study period.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by participants. All analyses will be run on the AMS, unless specified otherwise.

- Summary of ADAb status overall and by each visit separated by treatment group
- Summary of the time-point of the first occurrence of ADAb positivity during the treatment period by treatment group. A plot of the titer by time to first ADAb positivity will be prepared.
- All individual participant-level ADAb results will be listed.
- The number and percentage of participants in each of the 8 ADAb categories during the treatment period by treatment group.
- The prevalence of immunogenicity, separated by treatment group, and defined subcategory, will be reported by visit, defined as (cumulative) proportion of participants having confirmed positive ADAb samples at any visit up to and including that visit. Missing samples will not be included in the denominator.
- The time to achieving treatment-emergent ADAb positivity, separated by treatment group and defined subcategory, will be analyzed based on Kaplan-Meier methods. This will be shown only for Categories 2 and 8 above. Participants will be considered to have an event at the time point at which treatment emergent ADAb positive is first achieved (taking the MSR into consideration for sub-category 5). Participants classified as treatment-emergent ADAb negative will be censored at the time of the last available ADAb result.
- A summary of HiSCR₅₀ responders at Week 16, separated by treatment group, as a function of ADAb titer will be presented graphically.
- Individual plots of plasma bimekizumab concentrations/ ADAb titer both plotted on the Y-axes by visit (x-axis) for the full treatment period (excluding SFU for interim analyses and

including SFU for final analyses) will be presented for participants with and without HiSCR₅₀ response at Week 16.

- Spaghetti plots of ADA_b titer (y-axis) by visit (x-axis), separated by treatment group for all ADA_b positive participants, including Baseline positive participants.
- Box plots of ADA_b titer (logscale) by time to first ADA_b positivity by treatment group.

The groups for defining ADA_b status for safety subgroup analyses are as follows:

- AEs starting before first ADA_b positive result
- AEs starting on or after first ADA_b positive result
- AEs for participants who were always ADA_b negative

This is further explained in Section [10.2.2](#)

10 SAFETY ANALYSES

All analysis of safety variables will be performed using the SS, MS, and AMS.

The AMS will be used for summaries of safety that include data from the Initial Treatment Period and Maintenance Treatment Period.

Summaries of safety will be presented for the Initial Treatment Period, Maintenance Treatment Period, and combined Initial and Maintenance Treatment Period unless specified otherwise.

10.1 Extent of exposure

Summaries for exposure will be provided. This consists of a descriptive summary of study medication duration in days. In addition, total study medication duration and time at risk will be summarized in years by treatment group and treatment period (ie, the Initial Treatment Period, the Maintenance Treatment Period, and the Initial and Maintenance Treatment Period). Summary of exposure in Maintenance Treatment Period will be on MS. The cumulative study medication duration will be summarized for study participants exposed for given durations of time. For the cumulative duration through Week 48 the following categories for duration will be used:

- >0 weeks
- ≥4 weeks
- ≥8 weeks
- ≥12 weeks
- ≥16 weeks
- ≥20 weeks
- ≥24 weeks
- ≥28 weeks
- ≥32 weeks
- ≥40 weeks
- ≥48 weeks

Definitions for study medication duration and time at risk in days are provided below for each period. Time at risk will be summarized in years. Time at risk in years is calculated by dividing the time at risk in days by 365.25.

Throughout this section, date of last clinical contact for each participant is defined as the maximum of (last visit date including SFU visit, last imputed AE start date, date of study termination or completion, last date of study drug administration).

10.1.1 Exposure during the Initial Treatment Period

Definitions for study medication duration (days) and time at risk (days) during the Initial Treatment Period are provided as follows:

10.1.1.1 Study medication duration (days)

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Period + 14 days.

Note: The use of 14 days assumes a Q2W dosing interval (bimekizumab 320mg Q2W and placebo). For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Initial Treatment Period – Date of first dose in the Initial Period + 28 days).

Note: If date of last dose in the Initial Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Treatment Period + 1.
- For participants who die during the Initial Period, if date of last dose in the Initial Period + 14 days (or date of last bimekizumab dose in the Initial Treatment Period + 28 days in the case of Q4W dosing) extends to a date beyond the date of death, then this calculation reverts to:
 - Date of death – Date of first dose in the Initial Period + 1.

10.1.1.2 Time at risk (days)

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Week 16 visit and continue to the Maintenance Treatment Period:
 - Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Period + 1.
- For participants who discontinue on or prior to the final visit of the Initial Period, use the minimum of the following:
 - Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Treatment Period + 141
 - The total number of days in the Initial Treatment Period (112 days). For AEs that emerged after 112 days but still within the 140 days window, those AEs would be

classified as TEAE, but will be excluded from the output based on the Initial Treatment Period. However, these AEs will be included in the AE summaries for Maintenance Treatment Period.

- Date of last clinical contact – Date of first dose in the Initial Treatment Period + 1.
- For participants who die prior to the final visit of the Initial Treatment Period: Date of death – date of first dose in the Initial Period + 1.

10.1.2 Exposure during the Maintenance Treatment Period

Definitions for study medication duration (days) and time at risk (days) during the Maintenance Treatment Period are provided as follows:

10.1.2.1 Study medication duration (days)

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 14 days.

The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W in the Maintenance Treatment Period, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 28 days).

Note: If date of last dose in the Maintenance Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the final visit date of the Maintenance Treatment Period (not including SFU), then this calculation reverts to:

- Final visit date of the Maintenance Treatment Period (not including SFU) – date of first dose in the Maintenance Treatment Period + 1.
- For participants who die during the Maintenance Treatment Period, then this calculation reverts to:
 - Date of death – Date of first dose in the Maintenance Treatment Period + 1.

10.1.2.2 Time at risk (days)

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study):
 - Date of last visit of the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + 1.
- For participants who die prior to the final visit of the Maintenance Treatment Period:
 - Date of death – Date of first dose in the Maintenance Period + 1.
- For all other participants, use the minimum of the following:
 - Date of last dose in the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + 141 days.

- Date of last clinical contact – Date of first dose in the Maintenance Treatment Period + 1.

Note: This group could include participants who discontinue the Maintenance Treatment Period early, participants who complete the Maintenance Treatment Period as scheduled but choose not to continue into an extension study, or participants who are ongoing in the SFU period at the time of the data snapshot.

10.1.3 Exposure during the Initial and Maintenance Treatment Period

Definitions for study medication duration (days) and time at risk (days) during the Initial and the Maintenance Treatment Period are provided as follows:

10.1.3.1 Study medication duration (days)

Definitions for study medication duration (days) are provided as follows:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Period.

Note: The algorithms for calculating these durations are specified in Section 10.1.1.1 and Section 10.1.2.1.

Note: If date of last dose in the Initial Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Periods - 1.

10.1.3.2 Time at risk (days)

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study):
 - Final visit date – Date of first dose + 1.
- For participants who die prior to the final visit:
 - Date of death – Date of first dose in the + 1.
- For all other participants, use the minimum of the following:
 - Date of last dose – Date of first dose + 141 days.
 - Date of last clinical contact – Date of first dose + 1.

Note: This group could include participants who discontinue early, participants who complete the Maintenance Treatment Period as scheduled but choose not to continue into an extension study, or participants who are ongoing in the SFU period at the time of the data snapshot (in the case of the interim analysis).

10.2 Adverse events

An AE is any untoward medical occurrence in a patient or clinical study participant, temporally associated with the use of IMP, whether or not considered related to the IMP. An AE can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease (new or exacerbated) temporally associated with the use of IMP.

10.2.1 Data considerations

Treatment-emergent AEs are defined as those AEs that have a start date on or following the first dose of study treatment through the final dose of study treatment + 140 days (covering the 20-week SFU period). If it is not possible (due to partial dates) to determine whether or not an AE is treatment-emergent then it will be assumed to be a TEAE.

An AE will be assigned to the Initial Treatment Period if it started between the first administration of IMP on Day 1 up to Week 16. An AE will be assigned to the Maintenance Treatment Period if it started between the Week 16 study drug administration and Week 48.

If an AE occurs on the date of a treatment switch, the event is attributed to the original treatment. The only exception to this is if the AE fulfills any of the criteria specified below:

- Events that fulfill the anaphylaxis criteria for acute events (refer to Appendix A in Section 12.1)
- Hypersensitivity events identified by the SMQ “Hypersensitivity (SMQ)” (see Section 12.1 Appendix A)
- Events with an high level term (HLT) of “Administration site reactions NEC”
- Events with an HLT of “Injection site reactions”

The rules for imputing partial start or stop dates are outlined in Section 4.2.4.

Any TEAEs that occur during the SFU Period will be attributed to the period in which the participant was before initiating the SFU Period.

Duration of AEs will not be calculated if there is missing stop date information.

If the intensity of an AE is unknown, it is considered as severe. If the relationship to study drug is missing, it is considered as related. Note that if the seriousness of an adverse event is unknown, every attempt should be made to resolve this prior to a snapshot for an interim analysis or database lock; in the exceptional case that the seriousness of an adverse event is still missing then no imputation should be applied for this characteristic.

Adverse events will be presented as “number of participants (percentage of participants) [number of events]”. In this style of output, “[number of events]” will include all cases of an AE including repeat occurrences in individual participants, while “number of participants” will count each participant only once.

Subject time at risk represents the time a participant is at risk for having an AE. The definitions for subject time at risk (in days) are outlined in Section 10.1. These definitions will be used for exposure-adjusted AE summaries.

Selected AE summaries will include the exposure-adjusted incident rate (EAIR) with associated 95% CI and the exposure adjusted event rate (EAER).

The EAIR is defined as the number of participants (n) with a specific AE adjusted for the exposure and will be scaled to 100 subject-years:

$$EAIR = 100 \times n / \sum_{i=1}^N (T_{Exp(i)})$$

Where $T_{Exp(i)}$ is the exposure time and N is the number of participants at risk.

If a participant has multiple events, the time of exposure is calculated to the first occurrence of the AE being considered. If a participant has no events, the total time at risk is used.

Exact Poisson 95% confidence intervals for incidence rates are calculated using the relationship between the Poisson and the Chi-square distribution (Ulm, 1990; Fay and Feuer, 1997):

$$LCL = \chi^2_{2n, \alpha/2} / 2$$

$$UCL = \chi^2_{2(n+1), 1-\alpha/2} / 2$$

where n is the number of participants with a specific AE for the incidence rate of interest and is the basis for the number of the degrees of freedom for the chi-square quantile for the upper tail probability χ^2 .

The EAER will be the number of AEs including repeat occurrences in individual participants divided by the total time at risk scaled to 100 patient-years and calculated using:

$$EAER = 100 \times N_{AE} / \sum_{i=1}^N (T_{Risk(i)})$$

where N_{AE} is the total number of AEs, T_{Risk} is the time at risk for each participant, and N is the total number of participants at risk.

No confidence interval will be computed for EAER.

Selected summaries, as specified in Section 10.2.2, will include the risk difference between bimekizumab and placebo. The risk difference is calculated as:

$$RD = IP_{BKZ} - IP_{PBO}$$

where IP_{BKZ} is the incidence proportion for the bimekizumab-treated group and IP_{PBO} is the incidence proportion for the placebo group. Note that incidence proportion simply refers to the percentage of participants within the specified treatment group that experienced a given adverse event.

The standard error for the risk difference is calculated as follows:

$$SE_{RD} = \sqrt{\left(IP_{BKZ} \times \left(\frac{1 - IP_{BKZ}}{n_{BKZ}} \right) \right) + \left(IP_{PBO} \times \left(\frac{1 - IP_{PBO}}{n_{PBO}} \right) \right)}$$

where n_{BKZ} is the number of participants in the bimekizumab-treated group and n_{PBO} is the number of participants in the placebo group.

The corresponding confidence interval for the risk difference is as follows:

$$CI_{RD} = RD \pm Z_{1-\alpha/2} \times (SE_{RD})$$

where $Z_{1-\alpha/2}$ is the Z statistic for the corresponding level of alpha. For the risk difference confidence intervals calculated in this SAP, 1.96 will be used (corresponding to a two-sided alpha of 0.05 and 95% confidence interval). The risk difference and corresponding CI will be displayed as percentage.

10.2.1.1 COVID-19 related considerations

To assess the impact of COVID-19 mass vaccination on TEAEs, a sensitivity analysis will present all TEAEs excluding TEAEs assessed as exclusively related to COVID-19 vaccine by the investigator. TEAEs recorded as related to both study medication and COVID-19 vaccination should not be excluded. A complementary table and listing of TEAEs related to COVID-19 vaccine will be presented.

Another sensitivity analysis will present all TEAEs excluding TEAEs with start date on or up to 5 days after date of COVID-19 vaccine. Note that study participants may receive more than one administration of COVID-19 vaccine. A complementary table and listing of TEAEs with start date on or up to 5 days after date of COVID-19 vaccine will also be presented.

10.2.2 AE summaries

The following summaries will be provided by treatment group for the Initial Treatment Period, and the Initial and Maintenance Treatment Period combined based on the SS, and AMS respectively. In addition, all summaries of TEAEs based on “100 subject years” will include EAIR (with 95% confidence interval) and EAER. For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE. These AEs will be excluded from the outputs based on the Initial Period but included in the AE summaries for Initial and Maintenance Treatment Period.

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Study Participant Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT
- Incidence of TEAEs by Maximum Relationship by SOC, HLT, and PT
- Incidence of Serious TEAEs by Relationship SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of Related Serious TEAE by SOC, HLT, and PT
- Incidence of Severe TEAE per 100 subject years by SOC, HLT, and PT

- Incidence of TEAEs Leading to Death by Relationship by SOC, HLT, and PT – Note: For EudraCT reporting purposes
- Incidence of TEAEs by Maximum Severity, SOC, HLT, and PT
- Incidence of TEAEs by decreasing frequency of PT
- Incidence of TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs by SOC, HLT, and PT
- Incidence of Non-Serious TEAEs by Maximum Relationship SOC, HLT, and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of Non-Serious TEAEs Above Reporting Threshold of 5% by Relationship SOC and PT
- Incidence of Related TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Related TEAEs Above Reporting Threshold of 5% by SOC and PT
- Incidence of TEAEs – Suspected and Confirmed COVID-19 cases by SOC, HLT and PT
- Incidence of TEAEs Excluding TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval Censored TEAEs by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval TEAEs by SOC, HLT, and PT

Suspected and confirmed COVID-19 cases will be identified with the preferred terms “Corona virus infection” or “Corona virus test positive”.

The following subset of tables will also be presented for the Maintenance Treatment Period using the MS:

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT

The following tables will be presented for the Initial Treatment Period:

- Incidence of TEAEs Above Reporting Threshold of 5% with Risk Differences by SOC and PT
- Incidence of Serious TEAEs and Risk Differences by SOC and PT

The following table will be presented for the combined Initial and Maintenance Treatment Period. This summary will include only AEs that occur while a participant is on bimekizumab.

Any AEs in the Initial Treatment Period that begin while a participant is on placebo will be excluded.

- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status. This will include columns for the following:
 - TEAEs starting before the first ADA b positive result (includes ADA b categories 2 and 5) where TEAEs have occurred before the following events: a) the first positive ADA b result for subjects in category 2 and b) the first post-Baseline boosted ADA b titer result for subjects with titer results and the first post-Baseline positive ADA b result for subjects with positive ADA b at Baseline with no other samples with titer available for subjects in category 5
 - TEAEs starting on the same date or after the first ADA b positive result (includes ADA b Categories 2, 3, 4, 5 and 6) where TEAEs have occurred on or after the following events: a) the first positive ADA b results for subjects in categories 2, 3, 4 and 6, and b) the first post-Baseline boosted ADA b titer result for subjects with titer results and the first post-Baseline positive ADA b result for subjects with positive ADA b at Baseline with no other samples with titer available for subjects in category 5
 - TEAEs for subjects who are ADA b negative at all timepoints (includes ADA b Category 1)

The tables with risk differences will also be accompanied by figures (dot plots) which show the incidence of the adverse events and corresponding 95% risk difference confidence intervals. These will be ordered by descending order of risk difference (bimekizumab vs placebo).

10.2.3 Other Safety topics of interest

The following are AEs considered to be other safety topics of interest that require special statistical analyses. Along with the tables described, there will be a table which displays the risk difference and 95% confidence intervals for each of the topics of interest in the Initial Treatment Period. A corresponding figure (with dot plots) will be prepared.

A by-participant listing of all AEs of safety topics of interest will be presented by type of safety topics of interest.

10.2.3.1 Infections (serious, opportunistic, fungal and TB)

- **Incidence of Serious Infection TEAEs per 100 subject years by SOC, HLT and PT**

Serious infections will be identified based on MedDRA classification (SOC “Infections and infestations”) using the “Any SAE” table. A separate table does not need to be produced to summarize these events.

- **Incidence of Fungal Infection TEAEs per 100 subject years by SOC, HLT and PT**

Fungal infections will be summarized in a stand-alone table. The table will include all TEAEs (serious and non-serious) that code into the High Level Group Term (HLGT) “Fungal infectious disorders”

- **Incidence of Opportunistic Infection TEAEs per 100 subject years by SOC, HLT and PT**

Opportunistic infections (including tuberculosis) will be summarized in a stand-alone table. The table will include all TEAEs identified using UCB-defined search criteria .

The following steps will be followed for identifying and reviewing opportunistic infections:

Identification Process

The steps below outline 2 ways in which opportunistic infections (or potential opportunistic infections) can be identified:

Step 1: Refer to column B of the spreadsheet, which identifies the PTs to be classified as opportunistic infections using either a single 'x' or a double 'xx'.

- All TEAEs that code to a PT flagged with a single 'x' need to also be serious to be considered an opportunistic infection.
- All TEAEs that code to a PT flagged with a double 'xx' are considered to be an opportunistic infection, regardless of seriousness.

Step 2: Refer to column C of the spreadsheet, which identifies the PTs that need to be evaluated on a case-by-case basis by the study physician to determine whether or not it is an opportunistic infection. If column C has a single 'x', then the corresponding preferred term will be flagged for case-by-case review by the study physician.

Review Process

Opportunistic infections for a given study will be reviewed on the following occasions:

- At quarterly Infectious Disease Committee (IDC) Meetings, listings will be produced for each study (see details below) and reviewed by the corresponding study physician ahead of the IDC Meeting. These listings will be posted as part of the broader Safety Signal Detection (SSD) deliverable to a folder named for the given quarter (eg, 2018Q4) on the SharePoint. They will be based on the same data cut as the one used for SSD and will be delivered at the same time as the SSD outputs. The IDC will then agree on the final adjudication for each potential opportunistic infection.
- For each study, a final listing for opportunistic infections (in the format described below) will be produced and agreed upon between the study physician and the IDC prior to finalizing the database.

In each of the circumstances described above, the study programming team will produce an Excel listing that will be provided to the project lead statistician, project lead programmer, and to the study physician (who will subsequently provide it to the IDC). The Excel listing will contain the following columns (using the descriptions below as the column headings in the Excel listing):

- Study ID
- Unique Participant ID
- AE Term (Verbatim)
- AE Preferred Term
- AE System Organ Class
- AE High Level Term

- AE Low Level Term
- Date of Onset
- Outcome of Adverse Event
- Date of Outcome
- TEAE Flag
- Serious Adverse Event?
- Relationship to Study Medication
- Intensity
- Action Taken with IMP
- Opportunistic Infection – Automatic
- Opportunistic Infection – Manual Review
- Flag
- Data Cut Date
- Opportunistic Infection – Final Adjudication

Note the following about the final 5 variables in this listing:

- *Opportunistic Infection – Automatic*: This is flagged as “Y” if the criteria for automatic selection as described in “Step 1” of the identification process are met.
- *Opportunistic Infection – Manual Review*: This is flagged as “Y” if the criteria for case-by-case selection as described in “Step 2” of the identification process are met.
- *Flag* – This has a value of either “NEW” or “OLD”. It is marked as “NEW” if the event is appearing for the first time in that run of the listing. Otherwise, if it has appeared previously, it is marked as “OLD”. Unique records are determined by USUBJID AESPID for purposes of identifying whether an event has been modified from a previous run.
- *Date* – Only for cases where Flag is “NEW”, this field will be populated with the data cut date for that particular run of the listing.
- *Opportunistic Infection – Final Adjudication* – For new events, this is always left blank by the programmers. It will be completed by the study physician/IDC for every event that appears in the listing. For events adjudicated as opportunistic, the field will be populated with a “Y”.

Following each review by the study physician and IDC, the Opportunistic Infection – Final Adjudication column will be completed (as described above), and the spreadsheets for each study will be returned to the study programming team via e-mail (coordinated by the IDC secretary). Then, for subsequent runs of the listing, the study programming teams will incorporate adjudications from previous runs.

10.2.3.2 Malignancies

- **Incidence of Malignant or Unspecified Tumours TEAEs per 100 subject years by SOC, HLT and PT**

These events will be presented in the following tables:

- One table will be based on the criteria Standardized MedDRA Query (SMQ) = “Malignant or unspecified tumours (SMQ)”
- One table will be based on the criteria SMQ = “Malignant tumours (SMQ)”.

SMQ search will include all TEAEs which code to a PT included in the Scope=Narrow group within each SMQ.

Note that the events included in the “Malignancies” table will be a subset of the events included in the “Malignancies (including unspecified)” table. While the “Malignant tumours (SMQ)” is most relevant, “Malignant or unspecified tumours (SMQ)” must be reviewed for potential malignancies.

The output tables will include 2 different overall incidence rows:

- The first overall incidence row will summarize “Any Malignancy” and this row will summarize the incidence of all AEs flagged for inclusion in the table (using the appropriate SMQ depending on the table), regardless of the HLT it codes to.
- The second overall incidence row will summarize “Any Malignancy excluding non melanomic skin cancers HLT” and this row will summarize the incidence of AEs flagged for inclusion in the table (using the appropriate SMQ depending on the table), excluding those which code to an HLT of “skin neoplasms malignant and unspecified (excl melanoma)”.

10.2.3.3 Major adverse cardiac event

- **Incidence of Adjudicated Major Cardiac Event (MACE) TEAEs per 100 subject years by SOC, HLT and PT**

Potential cardiovascular events are adjudicated by the independent Cardiovascular Event Adjudication Committee (CV-CAC) according to the CV-CAC Charter (version 6.0).

Adjudicated events are classified by the CV-CAC to one of the event types as defined in [Table 10–1](#). The classification of an event as a Major Adverse Cardiac Event (MACE) is also determined by the CV-CAC. Events which are classified by the CV-CAC as any of the event types identified in the third column of [Table 10–1](#) will be considered an extended MACE. Note that extended MACE is determined programmatically and includes a broader scope definition of MACE.

MACE as determined by the CV-CAC will be presented in a table and listing. Extended MACE will be presented separately in a table and listing.

Another table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be a MACE and/or an extended MACE.

Additionally, a listing of all events identified for potential review by the CV-CAC will be produced. This listing will indicate whether each event was identified by the CV-CAC Chair for full committee review.

Table 10–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
1	Non-Fatal Myocardial Infarction (MI)	Yes
2	Non-Fatal Stroke: hemorrhagic	Yes
3	Non-Fatal Stroke: ischemic	Yes
4	Non-Fatal Stroke: embolic	Yes
5	Non-Fatal Stroke: undeterminable	Yes
6	Hospitalization or ER for Unstable Angina with urgent revascularization	Yes
7	Hospitalization or ER for Unstable Angina without urgent revascularization	No
8	Hospitalization for Heart Failure	Yes
9	Transient Ischemic Attack (TIA)	No
10	Coronary Revascularization Procedures (e.g. percutaneous coronary intervention, coronary artery bypass grafting)	Yes
11	Urgent Revascularization Procedures (i.e. due to symptoms of brain ischemia or pending infarction)	Yes
12	Arrhythmia (not associated with ischemia)	No
13	Peripheral Arterial Event	No
14	Venous Thromboembolic Event: pulmonary embolism (PE)	No
15	Venous Thromboembolic Event: deep vein thrombosis (DVT)	No
16	Venous Thromboembolic Event: PE and DVT	No
17	Other CV Event	No
18	Death due to Myocardial Infarction (MI)	Yes
19	Death due to Stroke	Yes
20	Sudden Cardiac Death	Yes

Table 10–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
21	Other CV Death (e.g. heart failure, pulmonary embolism, cardiovascular procedure-related)	Yes
22	Cardiovascular: Undetermined Cause of Death (i.e. cause of death unknown)	Yes
23	Non-Cardiovascular Death	No
24	Non-Cardiovascular Event	No
99	Inadequate information to adjudicate	No

CV=Cardiovascular; DVT=Deep Vein Thrombosis; ER=Emergency Room; MACE=Major Adverse Cardiac Event; MI=Myocardial Infarction; PE=Pulmonary Embolism; TIA=Transient Ischemic Attack.

MACE is determined by the adjudication committee and is not identified programmatically based on event type.

10.2.3.4 Neutropenia

• Incidence of Neutropenia TEAEs per 100 subject years by SOC, HLT and PT

This table will be based on the following PTs (regardless of seriousness):

- Autoimmune neutropenia
- Band neutrophil count decreased
- Cyclic neutropenia
- Febrile neutropenia
- Idiopathic neutropenia
- Neutropenia
- Neutropenic infection
- Neutropenic sepsis
- Neutrophil count decreased

10.2.3.5 Suicidal Ideation and Behavior

• Incidence of Suicidal Ideation or Behavior TEAEs per 100 subject years by SOC, HLT and PT

Potential neuropsychiatric events are adjudicated by the independent Neuropsychiatric Adjudication Committee according to the Neuropsychiatric Adjudication Committee (version 8.0). Adjudicated events are classified by the Committee as Suicidal or Non-suicidal. Adjudicated events are also further classified by the Committee to one of the event types as defined in [Table 10–2](#). Suicidal Ideation and Behavior (SIB) is defined as events classified by the Committee as Suicidal.

A table and listing will present SIB events.

Another table and listing will present the adjudicated neuropsychiatric events by type. For each neuropsychiatric event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be Suicidal or Non-Suicidal. For event type suicidal ideation, the listing will also indicate if intent was present and if the suicidal ideation was clinically significant.

Additionally, a listing of all events identified for potential review by the Committee will be produced. This listing will indicate whether each event was identified by the Neuropsychiatric Event Adjudication Committee Chair for full committee review.

Table 10–2: Neuropsychiatric event classifications

Event Type Code	Event Classification	Event Type
1	Suicidal	Suicidal events/completed suicide
2	Suicidal	Suicide attempt
3	Suicidal	Preparatory acts toward imminent suicidal behavior
4	Suicidal/Non-suicidal ^a	Suicidal ideation
7	Non-suicidal	Nonsuicidal Self-injurious behavior
8	Non-suicidal	Nonsuicidal Other
99	Not applicable	Inadequate information to adjudicate

^a Suicidal ideation event types can be classified by the Neuropsychiatric Adjudication Committee as Suicidal or Non-suicidal depending on whether intent to die was present.

10.2.3.6 Inflammatory bowel disease

- Incidence of Inflammatory Bowel Disease TEAEs per 100 subject years by SOC, HLT and PT**

Selected gastrointestinal events are adjudicated by the independent Inflammatory Bowel Disease (IBD) Adjudication Committee (IBD-CAC) according to the IBD-CAC Charter (version 3.0). Adjudicated events are classified by the IBD-CAC into one of the diagnostic types as defined in [Table 10–3](#). The events will further be classified as definite, probable or possible IBD.

An overview of adjudicated IBD events will be stratified by subjects with or without a previous medical history of IBD. Previous medical history of IBD will be determined using the information recorded on the History of IBD CRF page ("Does subject have a history of IBD?"). This overview table will present events adjudicated by the IBD-CAC as either possible, probable or definite IBD. Definite and probable IBD will also be aggregated and summarized in this table. In addition, this table will summarize each IBD event classification (possible, probable or definite) separately.

Another table and listing will present the adjudicated IBD events by type. For each IBD event type, the individual PTs which fall within each event type will be summarized.

Additionally, a listing of all events identified for potential review by the IBD-CAC will be produced. This listing will indicate whether each event was identified by the IBD-CAC Chair for full committee review.

A further supportive listing will present the individual diagnostic criteria met for each adjudicated IBD event.

Table 10–3: IBD event classifications

Event Type Code	Event Type (Classification and diagnosis)	Classification
1	Possible Inflammatory Bowel Disease – Crohn’s Disease	Possible
2	Probable Inflammatory Bowel Disease – Crohn’s Disease	Probable
3	Definite Inflammatory Bowel Disease – Crohn’s Disease	Definite
4	Possible Inflammatory Bowel Disease – Ulcerative Colitis	Possible
5	Probable Inflammatory Bowel Disease – Ulcerative Colitis	Probable
6	Definite Inflammatory Bowel Disease – Ulcerative Colitis	Definite
7	Possible Inflammatory Bowel Disease – type unclassified	Possible
8	Probable Inflammatory Bowel Disease – type unclassified	Probable
9	Definite Inflammatory Bowel Disease – type unclassified	Definite
10	Symptoms not consistent with Inflammatory Bowel Disease	Not applicable
11	Possible Inflammatory Bowel Disease – Microscopic Colitis	Possible
12	Probable Inflammatory Bowel Disease – Microscopic Colitis	Probable
13	Definite Inflammatory Bowel Disease – Microscopic Colitis	Definite
14	Possible Inflammatory Bowel Disease – no further differentiation possible	Possible
15	Probable Inflammatory Bowel Disease – no further differentiation possible	Probable
16	Definite Inflammatory Bowel Disease – no further differentiation possible	Definite
99	Not enough information to adjudicate	Not applicable

IBD=inflammatory bowel disease.

Note: IBD diagnoses of “microscopic colitis” and “no further differentiation possible” were added in an adjudication charter amendment, accounting for the event type numbering.

10.2.3.7 Hypersensitivity (including anaphylaxis)

- Incidence of Anaphylactic Reaction TEAEs per 100 subject years by SOC, HLT and PT

A separate table will be prepared based on the MedDRA anaphylaxis Algorithm (see Appendix 1) for acute anaphylactic events (reported on the same day as when an injection was administered or 1 day after). An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

A separate table will be prepared to summarize hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. In addition, a separate table will be prepared to summarize serious hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All serious TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

Furthermore, a separate table will be prepared to summarize injection site reactions, identified using the HLTs: “Administration site reactions NEC” and “Injection site reactions”.

10.2.3.8 Hepatic events and PDILI

- **Incidence of hepatic events TEAEs per 100 subject years by SOC, HLT and PT**

A table for hepatic events will be created based on the SMQ of “Drug related hepatic disorders - comprehensive search (SMQ)”. However, these 2 sub-SMQs are to be excluded: “Liver neoplasms, benign (incl cysts and polyps) (SMQ)” and “Liver neoplasms, malignant and unspecified (SMQ)”. For each of the above SMQs, include all TEAEs which code to a PT included in the Scope=Broad and/or Scope=Narrow.

Note that all AEs meeting the above criteria are to be included. It will not be limited to events that the investigator determined to be related to study drug.

Cases of potential Hy’s Law will be reported separately in a liver function test table.

10.3 Clinical laboratory evaluations

Laboratory values, including markedly abnormal laboratory values will be presented descriptively by treatment group for the SS, MS, and AMS.

The markedly abnormal tables and those based on common terminology criteria for AEs (CTCAE) grade will be produced only for selected laboratory variables.

For tables where data are summarized by visit, unscheduled and repeat visits will not be summarized, but these data will be included in listings. For tables where multiple measurements over a period of time are considered (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered). All summaries will be presented in SI units and will be based on observed case values. In the case where laboratory values are below the LLQ, then these will be set to the midpoint between 0 and the lower limit of quantification for the purpose of summarizing the data. The following summaries are required:

- A summary of the absolute and change from Baseline values in each laboratory variable by treatment group and visit

- A summary of the number and percentage of participants experiencing markedly abnormal values at any time while on treatment (assessment on or following the first dose of study treatment through the minimum of period of interest (Week 16) or date of last dose + 140 days) by laboratory variable and treatment group. Two separate tables will show results for the Initial Treatment Period (for the SS) and the Initial and Maintenance Treatment Period (for the AMS).
- A summary of the number and percentage of participants with a given CTCAE grade (0,1,2,3, or 4) based on minimum/maximum post-baseline value by laboratory variable and treatment group. Two separate tables will show results for the Initial Treatment Period (for the SS) and the Initial and Maintenance Treatment Period (for the AMS).
- A shift table of the number and percentage of participants experiencing CTCAE grade 0,1,2,3, or 4 values (as applicable) at Baseline to maximum post-Baseline CTCAE grade, by laboratory variable and treatment group. Two separate tables will show results for the Initial Treatment Period (for the SS) and the Initial and Maintenance Treatment Period (for the AMS).
- A by-participant listing of all laboratory data (including urinalysis) will be provided. This listing will be presented by treatment group and will include: center, participant identifier, age, sex, race, weight, visit, laboratory variable, result (with abnormal values flagged as “L” or “H” accordingly) and unit.

Markedly abnormal values are defined as those with a severity of Grade 3 and above based on the CTCAE criteria Version 4.03. Definitions of markedly abnormal values using the Grade 3 cut points are given in the tables below for age ranges of ≥ 17 years [Table 10–4](#) for markedly abnormal liver function test values, [Table 10–5](#) for markedly abnormal biochemistry values and [Table 10–6](#) for markedly abnormal hematology values). Tables summarizing markedly abnormal values will include a summary (counts and percentages) of markedly abnormal labs observed at any time while on treatment (ie, treatment-emergent markedly abnormal [TEMA]). For this summary, Baseline values and values observed more than 140 days after the last administration of study medication are not considered. The laboratory results classified as Grade 3 or Grade 4 will be summarized and listed separately.

Table 10–4: Definitions of Markedly Abnormal Liver Function Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
ALP	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH
ALT	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH
AST	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH
Total Bilirubin	mg/dL	>3.0 x ULN	umol/L	>3.0 x ULN	AH
GGT	U/L	>5.0 x ULN	U/L	>5.0 x ULN	AH

Table 10–5: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Creatinine ¹	mg/dL	> 3.0 x Baseline or >3.0 x ULN	umol/L	> 3.0 x Baseline or >3.0 x ULN	AH
Glucose	mg/dL	<40 >250	mmol/L	<1.7 >13.9	AL AH
Calcium	mg/dL	>12.5 <7.0	mmol/L	>3.1 <1.75	AH AL
Magnesium	mg/dL	>3.0 <0.9	mmol/L	>1.23 <0.4	AH AL
Potassium	mmol/L	>6.0 <3.0	mmol/L	>6.0 <3.0	AH AL
Sodium	mmol/L	>155 <130	mmol/L	>155 <130	AH AL
Cholesterol	mg/dL	>400	mmol/L	>10.34	AH

¹ The markedly abnormal definitions for creatinine are based on the logical or, if either criterion is met the creatinine value will be designated as abnormal high.

Table 10–6: Definitions of Markedly Abnormal Hematology Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Hemoglobin	g/dL	<8.0 >4.0 above ULN	g/L	<80 >40 above ULN	AL AH
Lymphocytes Absolute	10 ⁹ /L	<0.5 >20.0	10 ⁹ /L	<0.5 >20.0	AL AH
Neutrophils Absolute	10 ⁹ /L	<1.0	10 ⁹ /L	<1.0	AL
Platelets	10 ⁹ /L	<50	10 ⁹ /L	<50	AL
WBC/Leukocytes	10 ⁹ /L	<2.0 >100	10 ⁹ /L	<2.0 >100	AL AH

Abbreviations: AH=abnormal high; AL=abnormal low; ALP = alkaline phosphatase; ALT= alanine aminotransferase; AST = aspartate aminotransferase; dL = deciliter; GGT: gamma-glutamyltransferase; L = liter; mg = milligram; mmol = millimoles; µg = microgram; ULN = upper limit of normal, WBC=white blood cells.

The table for markedly abnormal liver function tests (LFTs) will contain data beyond the CTCAE Grade 3 thresholds outlined in [Table 10-1](#) above in order to allow for a more thorough review of elevated LFTs. There will be 1 table, which will list the count and percentage of participants meeting the below criteria at any time during the study:

- AST: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- AST or ALT: >3xULN, >5xULN, >8xULN, >10xULN, >20xULN
- Total Bilirubin: >1.5xULN, >2xULN
- ALP: >1.5xULN

For any participant with at least one markedly abnormal LFT (AST >3xULN, ALT >3xULN, bilirubin >3xULN, or ALP >1.5xULN) the New Ratio (nR) will be calculated as the ratio of either ALT or AST (whichever is higher) to ALP, all expressed as multiples of their ULN as follows:

- $nR = [\text{maximum}(\text{AST}/\text{ULN or ALT}/\text{ULN})]/(\text{ALP}/\text{ULN})$

Any pDILI will be summarized (all criteria must be met at the same assessment):

- (AST or ALT > 3xULN) and Total Bilirubin > 1.5xULN
- (AST or ALT > 3xULN) and Total Bilirubin > 2xULN

In addition, a table will be produced to summarize potential Hy's Law cases. The following definition will be used in that table:

- $[\text{AST} \geq 3\text{xULN or ALT} \geq 3\text{xULN}] \text{ and Total Bilirubin} \geq 2\text{xULN}$ in the absence of ALP $\geq 2\text{xULN}$

In order to meet the above potential Hy's Law criteria, a participant must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation, if applicable) at the same assessment. For example, a participant who experiences a $\geq 2 \times \text{ULN}$ elevation of bilirubin at one visit and a $\geq 3\text{xULN}$ elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's Law criteria.

Potential hepatotoxicity (meeting one of the PDILI or Hy's Law laboratory criteria at least once) will be considered with and without symptoms potentially associated with hepatitis or hypersensitivity according to the investigator (reported on the Symptoms of Hepatitis and Hypersensitivity CRF page).

10.4 Vital signs, physical findings, and other observations related to safety

10.4.1 Vital signs

The following vital signs variables will be summarized: systolic blood pressure (mmHg), diastolic blood pressure (mmHg), body temperature (°C) and heart rate (beats/min). The following summaries will be provided for the SS:

- A summary of the absolute and change from Baseline value for each vital sign variable by treatment group and visit.
- A summary of the number and percentage of participants experiencing at least 1 markedly abnormal value for a vital sign variable as defined in Table 10–7, by treatment group and period (Initial Treatment Period [SS], and Initial and Maintenance Treatment Period [AMS]). Unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

Table 10–7: Definitions of Markedly Abnormal Blood Pressure Values

Parameter (unit)	Markedly Abnormal Low	Markedly Abnormal High
Systolic blood pressure (mmHg)	<90 and a decrease from Baseline of ≥ 20	>180 and an increase from Baseline of ≥ 20
Diastolic blood pressure (mmHg)	<50 and a decrease from Baseline of ≥ 15	>105 and an increase from Baseline of ≥ 15

A by-participant listing of all vital signs data will be provided. This listing will be presented by treatment group and will include: center, participant identifier, age, sex, race, weight, visit, vital sign variable and result (with abnormal values flagged as “L” or “H” accordingly).

10.4.2 Electrocardiograms

Electrocardiogram data will be analyzed by treatment group and visit for the SS.

A summary of the number and percentage of participants with normal, abnormal ECG results, as determined by the central reader, will be presented for all applicable visits.

The following ECG variables will be summarized (absolute values and change from Baseline) by visit: QT corrected for heart rate using Friderica’s formula (QTcF), RR, PR, QRS and QT.

QTc outliers are defined as QTcF values following dosing that are greater than 450 ms or are increases from Baseline greater than 30 ms. QTcF outliers will be highlighted in the data listing and summarized using the following categories:

- Values >450 ms, >480 ms, >500 ms
- Increase from Baseline of >30 ms, increase from Baseline of >60 ms, increase from Baseline of >90 ms
- Values >450 ms and increases of >30 ms. Values >500 ms and increases of >60 ms

The number and percentage of participant who meet the ECG outlier criteria at any assessment post first dose will be summarized for each period.

Two separate by-participant listing of all 12-lead ECG data will be provided based on interpretation from central reader and from site, respectively.

10.4.3 Other safety endpoints

For by-visit summaries, unscheduled and repeat visits will not be summarized, but these data will be included in listings. By-visit tables should include the SFU visit. Summaries over a period of

time (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

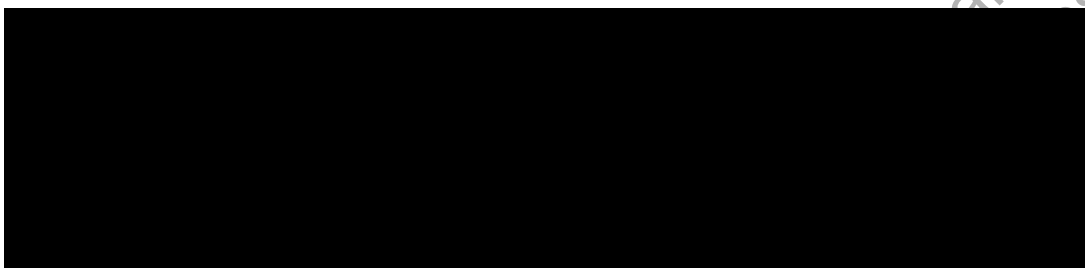
10.4.3.1 Physical examination

Abnormal results of the physical examination together with details of abnormalities: abnormality clinically significant or not, will be listed by participant and visit for SS.

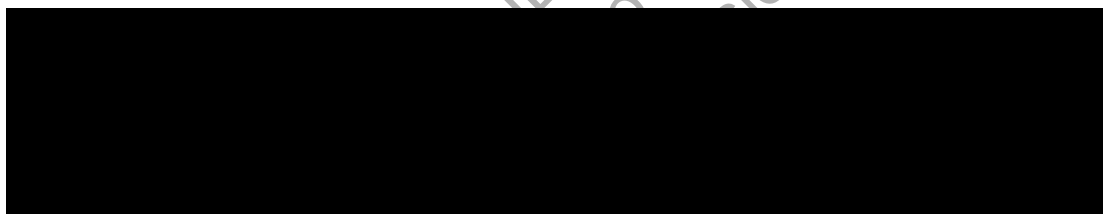
10.4.3.2 Columbia-Suicide Severity Rating Scale (C-SSRS)

The eC-SSRS questionnaire will be self-administered by the study participant and assessed by trained study personnel. This scale will be used to assess SIB that may occur during the study. Results of the eC-SSRS will be summarized using the number of participants and percentage with (i) suicidal ideation, (ii) suicidal behavior, (iii) suicidal ideation or behavior, and (iv) self-injurious behavior without suicidal intent.

Suicidal ideation is defined as an event in any of the following 5 categories:



Suicidal behavior is defined as an event in any of the following 4 categories:



Suicidal behavior or ideation is defined as an event in any of the above 9 categories.

Self-injurious behavior without suicidal intent is defined as an event in the category non-suicidal self-injurious injuries.

The incidence of participants with suicidal ideation, suicidal behavior, suicidal ideation or behavior, and self-injurious behavior will be summarized for the Initial Treatment Period and the combined Initial and Maintenance Treatment Period by treatment group.

A by-participant listing of the eC-SSRS questionnaire data will be provided by treatment group.

10.4.3.3 Assessment and management of TB and TB risk factors

A summary of the number and percentage of participants with negative, positive, and indeterminate IGRA (Interferon-Gamma Release Assay) results at Screening and Week 44 will be presented.

A by-participant listing of the “Evaluation of signs and symptoms of tuberculosis” questionnaire data and IGRA results will be provided by treatment group.

By-participant listing of the result of chest x-ray for tuberculosis will be provided by treatment group.

10.4.3.4 Pregnancy testing

Pregnancy testing will consist of serum testing at the Screening. The pregnancy test will be urine at all other visits.

A by-participant listing of the pregnancy test data will be provided by treatment group.

10.4.3.5 Patient Health Questionnaire (PHQ)-9 scores

The PHQ-9 is a multipurpose instrument for screening, diagnosing, monitoring, and measuring the severity of depression. The PHQ-9 scores for depression range from 0 to 27 with higher scores indicating worse state. A score of 5 to 9 is considered to be minimal symptoms of depression. If any of the 9 questions are missing, then the score is treated as missing. A score of 10 to 14 is considered minor depression, dysthymia, or mild major depression. A score of 15 to 19 is considered to indicate moderately severe major depression, and a score ≥ 20 is considered to be severe major depression.

Change from Baseline in PHQ-9 is derived as post-Baseline score minus Baseline score.

A summary of the absolute and change from Baseline value will be presented by treatment group and visit.

The percentage of study participants with scores below 5, between 5 and 9, between 10 and 14, between 15 and 19, and greater than or equal to 20 in PHQ-9 will be summarized as a shift from Baseline by visit and treatment group based on observed values.

The percentage of study participants with scores ≥ 15 at any post-Baseline visit and the number and percentage of study participants with scores ≥ 20 at any post-Baseline visit will be summarized by treatment group based on observed values. This summary will also include the percentage of study participants with increase from baseline ≥ 5 at any post-Baseline visit.

The number and percentage of participants that complete the PHQ-9 will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit.

11 REFERENCES

Common Terminology Criteria for Adverse Events (CTCAE); Version 4.0 June 2010.U.S. Department of Health and Human Services

Dauden E, Griffiths CE, Ortonne JP, Kragballe K, Molta CT, Robertson D, et al. Improvements in patient reported outcomes in moderate to severe psoriasis patients receiving continuous or paused etanercept treatment over 54 weeks: the CRYSTEL study. *J Eur Acad Dermatol Venereol*. 2009;23(12):1374-82.

Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997; 35(11):1095-108.

Esmann S, Dufour DN, Jemec GB. Questionnaire based diagnosis of hidradenitis suppurativa: specificity, sensitivity and positive predictive value of specific diagnostic questions. *Br J Dermatol*. 2010;163(1):102-6.

Guidance for Industry: E9 Statistical Principles for Clinical Trials. US Department of Health and Human Services, Food and Drug Administration. ICH 1998.

Hurley H. Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In: *Dermatologic surgery*. Roenigh R, Roenigh H, editors. Marcel Dekker. 1989:729-39.

Kimball AB, Kerdel F, Adams D, Mrowietz U, Gelfand JM, Gniadecki R. Adalimumab for the treatment of moderate to severe Hidradenitis suppurativa: a parallel randomized trial. *Ann Intern Med*. 2012;157(12):846-55.

Kimball AB, Jemec GB, Yang M, Kageliery A, Signorovitch JE, Okun MM, et al. Assessing the validity, responsiveness and meaningfulness of the Hidradenitis Suppurativa Clinical Response (HiSCR) as the clinical endpoint for hidradenitis suppurativa treatment. *Br J Dermatol*. 2014;171(6):1434-42.

Kimball AB, Okun MM, Williams DA, Gottlieb AB, Papp KA, Zouboulis CC, et al. Two Phase 3 Trials of Adalimumab for Hidradenitis Suppurativa. *N Engl J Med*. 2016a;375(5):422-34.

Kimball AB, Sobell JM, Zouboulis CC, Gu Y, Williams DA, Sundaram M, et al. HiSCR (Hidradenitis Suppurativa Clinical Response): a novel clinical endpoint to evaluate therapeutic outcomes in patients with hidradenitis suppurativa from the placebo controlled portion of a phase 2 adalimumab study. *J Eur Acad Dermatol Venereol*. 2016b;30(6):989-94.

Mallinckrodt CH, Lin Q, Lipkovich I, and Molenberghs G. A structured approach to choosing estimands and estimators in longitudinal clinical trials. *Pharmaceutical Statistics* 2012; 11:456-61.

O'Kelly, Michael, and Bohdana Ratitch. *Clinical trials with missing data: A guide for practitioners*. John Wiley & Sons, 2014

Ratitch, B., Lipkovich, I., O'Kelly, M. Combining Analysis Results from Mutiply Imputed Categorical Data, PharmaSUG, 2013, SP03

Rubin D.B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley and Sons.

Sartorius K, Emtestam L, Jemec GB, Lapins J. Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *Br J Dermatol*. 2009;161(4):831-9.

Sun H, Snyder E, Koch GG. Statistical planning in confirmatory clinical trials with multiple treatment groups, multiple visits, and multiple endpoints. *J Biopharmaceut Stat*. 2018;28(1):189-211.

UpToDate. 2011. Modified Hidradenitis Suppurativa Score (HSS). [ONLINE] Available at: <http://cursoenarm.net/UPTODATE/contents/mobipreview.htm?36/33/37400>. [Accessed 27 November 2017].

Zouboulis CC, Kyrgidis A, Jemec GBE, Bechara FG, Giamarellos Bourboulis EJ, Ingram JR, et al. Development and validation of IHS4, a novel dynamic scoring system to assess hidradenitis suppurativa/acne inversa severity. *British Journal of Dermatology*. 2017 June 21.

12 APPENDICES

12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The SMQ *Anaphylactic reaction* consists of three parts:

- A **narrow search** containing PTs that represent core anaphylactic reaction terms (Category A – core anaphylactic reaction terms)

Anaphylactic reaction
Anaphylactic shock
Anaphylactic transfusion reaction
Anaphylactoid reaction
Anaphylactoid shock
Circulatory collapse
Dialysis membrane reaction
Kounis syndrome
Shock
Shock symptom
Type I hypersensitivity

- A **broad search** that contains additional terms that are added to those included in the narrow search. These additional terms are signs and symptoms possibly indicative of anaphylactic reaction and categorized in B, C or D

▪ Category B (Upper Airway/Respiratory Terms)

Acute respiratory failure	Nasal obstruction
Asthma	Oedema mouth
Bronchial oedema	Oropharyngeal spasm
Bronchospasm	Oropharyngeal swelling
Cardio-respiratory distress	Respiratory arrest
Chest discomfort	Respiratory distress
Choking	Respiratory failure
Choking sensation	Reversible airways obstruction
Circumoral oedema	Sensation of foreign body
Cough	Sneezing
Cyanosis	Stridor
Dyspnoea	Swollen tongue
Hyperventilation	Tachypnoea
Irregular breathing	Throat tightness
Laryngeal dyspnoea	Tongue oedema

Laryngeal oedema	Tracheal obstruction
Laryngospasm	Tracheal oedema
Laryngotracheal oedema	Upper airway obstruction
Mouth swelling	Wheezing

▪ **Category C (Angioedema/Urticaria/Pruritus/Flush terms)**

Allergic oedema	Oedema
Angioedema	Periorbital oedema
Erythema	Pruritus
Eye oedema	Pruritus allergic
Eye pruritus	Pruritus generalised
Eye swelling	Rash
Eyelid oedema	Rash erythematous
Face oedema	Rash generalised
Flushing	Rash pruritic
Generalised erythema	Skin swelling
Injection site urticaria	Swelling
Lip oedema	Swelling face
Lip swelling	Urticaria
Nodular rash	Urticaria papular
Ocular hyperaemia	

▪ **Category D (Cardiovascular/Hypotension terms)**

Blood pressure decreased
Blood pressure diastolic decreased
Blood pressure systolic decreased
Cardiac arrest
Cardio-respiratory arrest
Cardiovascular insufficiency
Diastolic hypotension
Hypotension

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after, and for scenarios where two events must have been reported, both events must have occurred within one day of each other (as anaphylaxis is an acute event, imputed dates should not be used in the algorithmic approach):

- A narrow term or a term from Category A;
- A term from Category B - (Upper Airway/Respiratory) AND a term from Category C - (Angioedema/Urticaria/Pruritus/Flush);
- A term from Category D - (Cardiovascular/Hypotension) AND [a term from Category B - (Upper Airway/Respiratory) OR a term from Category C - (Angioedema/Urticaria/Pruritus/Flush)]

12.2 Appendix B: Definition of CTCAE grades

Table 12–1: Definition of CTCAE grades by biochemistry parameters						
Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine ¹	High	umol/L	>1-1.5x Baseline or >ULN-1.5 x ULN	>1.5-3.0x Baseline or >1.5 – 3.0 x ULN	>3.0x Baseline or >3.0 – 6.0 x ULN	>6.0 x ULN
Sodium	Low	mmol/L	130-<LLN	N/A	120-<130	<120
Sodium	High	mmol/L	>ULN-150	>150-155	>155-160	>160
Potassium ²	Low	mmol/L	3.0-<LLN	3.0-<LLN	2.5-<3.0	<2.5
Potassium	High	mmol/L	>ULN-5.5	>5.5-6.0	>6.0-7.0	>7.0
Calcium	Low	mmol/L	2.0-<LLN	1.75-<2.0	1.5-<1.75	<1.5
Calcium	High	mmol/L	>ULN-2.9	>2.9-3.1	>3.1-3.4	>3.4
Magnesium	Low	mmol/L	0.5-<LLN	0.4-<0.5	0.3-<0.4	<0.3
Magnesium	High	mmol/L	>ULN-1.23	N/A	>1.23-3.30	>3.30
Cholesterol	High	mmol/L	>ULN-7.75	>7.75-10.34	>10.34- 12.82	>12.82

1 The CTCAE Grade definitions for creatinine are based on the logical or the highest applicable CTCAE grade should be assigned to a creatinine value.

2 The decreased potassium criterion of 3.0-<LLN is specified for both CTCAE Grade 1 and Grade 2; values meeting this criterion will be counted as Grade 2.

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin	Low	g/L	100-<LLN	80-<100	<80	N/A
Hemoglobin ¹	High	g/L	>0-20 above ULN or >0-20 above Baseline if Baseline is above ULN	>20-40 above ULN or >20-40 above Baseline if Baseline is above ULN	>40 above ULN or >40 above Baseline if Baseline is above ULN	N/A
Platelets	Low	10 ⁹ /L	75-<LLN	50-<75	25-<50	<25

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
WBC	Low	$10^9/L$	$3 < LLN$	$2 < 3$	$1 < 2$	< 1
WBC	High	$10^9/L$	N/A	N/A	> 100	N/A
Lymphocytes	Low	$10^9/L$	$0.8 < LLN$	$0.5 < 0.8$	$0.2 < 0.5$	< 0.2
Lymphocytes	High	$10^9/L$	N/A	$> 4-20$	> 20	N/A
Neutrophils	Low	$10^9/L$	$1.5 < LLN$	$1.0 < 1.5$	$0.5 < 1.0$	< 0.5

LLN=lower limit of normal; N/A=not applicable; ULN=upper limit of normal, WBC=white blood cells

1 The CTCAE Grade definitions to be applied are dependent on the Baseline hemoglobin value. If the baseline value is $> ULN$ then the criteria relative to Baseline is applicable, otherwise the criteria relative to ULN is applicable.

Note that participants who meet the decreased potassium criterion of $3.0 < LLN$, which is specified as the decreased potassium lab criterion for both CTCAE Grade 1 and Grade 2, will be counted as Grade 2.

13 AMENDMENT TO THE STATISTICAL ANALYSIS PLAN

13.1 Amendment 1

Rationale for the amendment

The main purposes of this amendment were:

- General update to analyses to align with protocol amendment 3.
- Procedural clarifications from discussions and feedback provided at meetings
- Update to align with the bimekizumab program standards and safety topics of interest

Modifications and changes

Global Changes

Typos and formatting were updated throughout the document.

Global changes:

The following changes were made throughout the SAP:

- References to Section 3.9 were added through the efficacy sections to clarify definition and handling of intercurrent events
- PRO endpoint terminology was updated to match protocol (eg, Worst Pain score in HSSDD instead of Worst Pain in HSSDD)

Specific changes

In addition to the global changes, the following specific changes have been made (formats as missing spaces or redundant spaces are not listed, typos):

Change #1

The following abbreviations have been added:

CFS	COVID-19 Free Set
COVID-19	coronavirus disease 2019
NI	Negative Immunodepletion
nR	New Ratio
NS	Negative Screen
pDILI	potential drug induced liver injury
PI	Positive Immunodepletion
PS	Positive Screen

Change #2

Section 1 Introduction

The protocols were updated:

The SAP is based on the Protocol Amendment **3.2, 9 February 2021** ~~16 December 2019~~ and the Japan-specific amendment **3.1.2.1, 11 February 2021** ~~19 December 2019~~.

Change #3

Section 2.2 Study endpoints

The following text was added:

The endpoints based on HS Symptom Daily Diary (HSSDD) and Hidradenitis Suppurativa Symptom Questionnaire (HSSQ) pain responses are based on the current definitions, which are continuous. It is anticipated that a responder (binary) endpoint will be defined for the HSSDD and HSSQ pain items as well as other symptom items prior to database lock and unblinding, based on separate ongoing, blinded, psychometric analyses aiming to determine threshold for within-patient clinically meaningful improvement.

The below HSSDD and HSSQ pain response endpoints and analyses will be adjusted accordingly in a future SAP amendment.

Change #4

Section 2.2.1.2 Secondary Efficacy endpoints

The secondary endpoints were updated:

The secondary efficacy endpoints are defined as:

- HiSCR₇₅ response (defined as at least a 75% reduction from Baseline in the total AN count with no increase from Baseline in abscess or draining tunnel count) at Week 16
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline) by Week 16
- **Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16**
- Absolute change from Baseline in Skin Pain score at Week 16, as assessed by the “worst pain” item (11-point numeric rating scale) in the HS Symptom Daily Diary (HSSDD)
- ~~Absolute change from Baseline in Dermatology Life Quality Index (DLQI) Total Score at Week 16~~

Change #5

Section 2.2.1.3 Other efficacy endpoints

The following text was added:

- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) **and Total score**

Change #6

Section 2.2.3.2 Other safety endpoints

The other safety topics of interest were updated:

- Other safety topics of interest: infections (serious, opportunistic, fungal, and TB), neutropenia, hypersensitivity (**including anaphylaxis**), suicidal ideation and behavior, ~~depression~~, major **adverse** cardiovascular events, **hepatic events and potential drug-induced liver injury (PDILI)**, ~~function test changes/enzyme elevations~~, malignancies, and inflammatory bowel disease.

Change #7

Section 2.3.1 Study description

The following sentence was deleted:

Enrollment of study participants currently using antibiotics will be capped at 30% of overall enrollment.

Change #8

Section 3.1 General presentation of summaries and analyses

The following text was added:

For PRO continuous variables, descriptive statistics will also include variable score, absolute and percentage changes from baseline, Q1 and Q3, 10th, and 90th percentiles.

If no participants have data at a given time point, then only n=0 will be presented. The other descriptive statistics will be left blank. If n < 3 then the n, minimum, and maximum only will be presented. The other descriptive statistics will be left blank. If n = 3 n, mean, median, minimum and maximum will be presented only. The other descriptive statistics will be left blank.

For categorical variables, the number and percentage of study participants in each category will be presented. Unless otherwise noted, the denominator for percentages will be based on the number of study participants included in the respective analysis set. Study participants with missing data will be accounted for using the following approaches:

- For summaries of demographics and Baseline characteristics: percentages will be based on all study participants in the analysis set and a “Missing” category (corresponding to study participants with missing data for the variable being summarized) will be included as the last row in the list of categories being summarized.
- For summaries of efficacy and safety endpoints, unless otherwise specified: percentages will be based only on those study participants with observed data for the variable being summarized. As the denominator may be different from the number of study participants in the analysis set being considered, the denominator will be displayed in the table. The general format for displaying this will be “n/Nsub (%)”.

Percentages will be presented to 1 decimal place. If the percentage is 100%, a decimal will not be presented. If the count is 0, the percentage will not be presented. Typically, the % sign will be presented in the column header, but not with each individual value.

For the purpose of the tabulations the lower and upper confidence limits for the percentages will be truncated at 0 and 100% respectively. **Confidence intervals (CIs) for the response rates in efficacy summaries based on nonresponder imputation (NRI) will be computed using the Wilson approximation.**

Change #9

Section 3.3 Definition of Baseline

The following text was added:

For randomized participants for whom no start date of treatment is available, the Baseline value will be considered as the last available value on or before the randomization date.

Change #10

Section 3.4 Protocol deviations

The following text was added:

Deviations related to the Coronavirus Disease 2019 (COVID-19) global pandemic are unavoidable deviations from the protocol due to confirmed COVID-19 infection, suspected COVID-19 infection, general circumstances around COVID-19 without infection or any other deviation from the protocol due to COVID-19. COVID-19 protocol deviations will also be reviewed separately as part of the ongoing data cleaning process.

Change #11

The following section was added:

COVID-19 Free Set (Section 3.5.9)

The COVID-19 Free Set (CFS) will consist of all study participants randomized into the study and who have no COVID-19 impact through Week 16. This analysis set will be used for sensitivity analysis of the primary efficacy endpoint.

Change #12

Section 3.8 Coding dictionaries

The following sentence was updated:

Medications will be coded according to the World Health Organization Drug Dictionary (WHODD) version **MAR2021 B3 or later**. Medical procedures will not be coded.

Change #13

The following section was added:

Definition of an intercurrent event (Section 3.9)

Handling of intercurrent events is one of the key elements for the analysis of efficacy endpoints.

An intercurrent event is defined as receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy (See Section 8.2.2).

Receipt of systemic antibiotic rescue medication is defined as initiating any systemic antibiotic on or after Baseline for any reason (including in response to an AE). The only exception to this rule is if a participant randomized to the antibiotic stratum on a tetracycline antibiotic interrupts their stable dose of tetracycline antibiotic during the study and subsequently restarts the same tetracycline antibiotic as confirmed using the coded preferred term. The restarted dose and frequency of the antibiotic must be the same or lower than the regimen prior to the interruption.

The dates of an intercurrent event are as follows:

- For receipt of systemic antibiotic rescue medication: start date of the antibiotic
- For discontinuation of study treatment due to an AE or lack of efficacy: Last study treatment date + 17 days. Note: study treatment discontinuation includes study discontinuation.

The choice of 17 days is intended to capture the interval between dosing and lesion assessments (14 days), as well as the visit window (3 days).

An additional sensitivity analysis will be conducted where missing data due to COVID-19 will be considered an intercurrent event and will be imputed as a nonresponse at that particular visit. This will be identified when there are missing data at a visit that has been impacted by COVID-19 according to the COVID-19 impact CRF page. The date of this intercurrent event will be the date of the impacted visit.

Change #14

Section 3.10 Changes to protocol-defined analyses

The following text was added:

The HiSQOL endpoint was clarified to show that there are only 3 domains: symptoms, psychosocial, activities and adaptations **and to add total score**.

Also, the following endpoint was added to the list of protocol endpoints not included in the analysis:

- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48

Change #15

Section 4.1 Adjustments for covariates

The following text was added:

If a participant is stratified in the incorrect stratum (ie, the stratum recorded in the Interactive voice or web Response System differs from the actual stratum the participant belongs to), the actual stratum will be used for the analysis.

Change #16

Section 4.2.1.4 Missing Data Overview and Summary

The following text was added:

In summary, the approaches listed below will be used in this study for handling missing data for efficacy endpoints as appropriate:

- **NRI: Participants** who have missing data at the timepoint of interest are treated as though they did not respond to the treatment. **This approach is also referred to as Composite Estimand (NRI).**
- Multiple Imputation (MI) – MCMC / Monotone Regression: Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using monotone regression.

- **MI-MCMC / Reference-based imputation:** Using multiple imputation methodology, intermittent missing data are imputed based on the MCMC method, and monotone missing data are imputed using an imputation model based on placebo (reference) data.
- **LOCF:** Post-Baseline missing data are imputed by carrying forward the last available observation (including Baseline).
- **Tipping point analyses:** Assumptions will be made about average outcomes among the subsets of **participants** who prematurely discontinued study treatment. Various “delta adjustments” will be made to the assumed responses among missing data in each treatment arm with varying degrees of plausibility in order to identify assumptions about the missing data under which the conclusions change (O’Kelly, 2014). Then, the plausibility of such assumptions is discussed.
- **Observed case (OC):** Missing data are not imputed. Only **participants** with available data who have not discontinued study treatment at the given timepoint are considered. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing. ~~For OC summaries, intercurrent events are not handled differently than other missing data.~~
- **Treatment policy strategy:** All available data observed at the time point of interest will be considered, regardless of the occurrence of intercurrent events. This means the analysis includes on- and off-treatment values collected after study participants prematurely discontinued study treatment but agreed to remain on the study and continued to attend visits and provide assessments at those visits as well as values from study participants who received rescue antibiotic medication. Those observed values will be analyzed according to the study participant’s randomized treatment. Study participants for whom efficacy data cannot be obtained at the week of interest, despite attempts to retain them in the study, will have their data imputed using MI – MCMC / monotone.

Table 4-1 was updated:

Table 4–1: Missing data handling approach by endpoint priority and type								
Endpoint Priority	Endpoint Type	Composite Estimand (NRI)	Modified Composite Estimand (MI)	MI (MCMC/ Reference-based)	Tipping Point	Treatment Policy	Hypothetical Estimand	OC
Primary	Responder	S ^a	P	S ^a	S	S ^a		S
Secondary included in the statistical testing procedure	Responder	S ^a	P					S
	Continuous						P	S
	Binary	X	X					X

Secondary not included in statistical testing procedure	Continuous						X	X
Other	Responder	X^d	X					X^d
	Continuous						X	X^b
	Ordinal						X^c	X^c

B=Backup method, LOCF=Last observation carried forward, MCMC=Markov Chain Monte Carlo, MI=multiple imputation, NRI=Nonresponder imputation, OC=Observed case, P=Primary method, S=Sensitivity method, **X=Method to be used (no priority designated).**

Note: Composite estimand (NRI) refers Backup method is only applicable when the primary method is unable to the approach in which data preceded by the intercurrent event of study treatment discontinuation converge due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are also imputed as nonresponse.

Note: Modified Composite Estimand (MI) refers to the approach in which data preceded by the intercurrent event of study treatment discontinuation due to AE or lack of efficacy or receipt of rescue antibiotic medication are imputed as nonresponse, and other missing data are imputed via a multiple imputation model.

Note: Hypothetical Estimand (MI) refers to the approach where outcomes for study participants without an intercurrent event of study treatment discontinuation are as observed, and outcomes for study participants challenges with the intercurrent event are imputed via a multiple imputation model.

^a Imputation method is applied on continuous data, and responder endpoint is derived from the continuous endpoint based on complete data set where applicable.

^b Required only for by-visit summaries of variables whose value at Week 16 is part of the hierarchical testing procedure.

^c For variables with multiple categories, data will be summarized as observed with an additional missing row to capture missing data at a given visit.

^e Participants with intercurrent events are imputed as nonresponders for all subsequent timepoints before the imputation method is applied for all other missing data.

^d NRI/OC sensitivity analysis will be performed only for HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, HiSCR₁₀₀ summaries.

^e The treatment policy estimand will use the same MI-MCMC/Monotone Regression defined for the primary analysis, with the exception that participants with intercurrent events will not be treated as nonresponders for all subsequent timepoints before the imputation method is applied for missing data.

Change #17

Section 4.2.2 Missing data algorithms for efficacy analyses

The section has been moved from Section 4.2.1.5 and split into Section 4.2.2.1, Section 4.2.2.2, and Section 4.2.2.3.

Change #18

Section 4.2.2.1 MI – MCMC/Monotone Regression

This section was updated to the following text:

In many cases, missing efficacy data due to study treatment discontinuation should be dependent on the observed efficacy scores, but independent of unobserved data. This would be consistent

with a missing at random (MAR) pattern of missingness. To investigate the efficacy results under the assumption of data being MAR, a multiple imputation method will be applied, as follows:

Binary endpoint

For a binary endpoint (eg, HiSCR₅₀), the procedure is as follows:

6. Create a data set, one for each treatment group of participants with observed values and those needing estimation by multiple imputation. For the imputation step, a distinction is made between non-monotone missing values (ie, intermittent missing values between completed assessments) and monotone missing values (ie, where all participants data are missing after a given time point).
- c. For the intermittent missing values, the missing values in each data set will be filled in using the MCMC method with multiple chain, monotone missing data imputing pattern, and non-informative prior for all parameters. Unless specified differently, the first 200 iterations will not be used (the “Burn-in” option). A total of 100 sets of imputations will be performed. The seed used for these imputations will be 762 and all other multiple imputation procedures described in this SAP will use this same seed as well. The resulting 100 imputed data sets will have a monotone missing pattern and will be imputed using a method for monotone missingness. Note Create a data set, one for each treatment group (note that a separate imputation procedure must be invoked in SAS for each treatment group as the seed cannot be set for by groups beyond the first when using a by statement), of participants with observed values and those needing estimation by multiple imputation. The intermittent missing lesion counts in PROC MI.
- d. each data set (ie, missing values for a given subject that has available data before and after the missing timepoint) will be filled in using the MCMC method, with a total of 100 sets of imputations being performed. The seed used for these imputations will be 762 (note that all other multiple imputation procedures described in this SAP related to MCMC/Monotone regression analyses will use this same seed as well). For monotone missing data (ie, where all participant data is missing after a given timepoint), monotone regression will then be used to impute missing data. A separate regression model is estimated for each variable with missing values (ie, measurement at each time point). Based on the resulting model, a new regression model is then drawn and is used to impute the missing values for the variable. Since the data set has a monotone missing data pattern, the process is repeated sequentially for variables with missing values. This procedure will be based on the 100 imputed datasets generated from sets of imputations already created using the MCMC procedure and method such that there will be performed by imputation. The SAS® PROC MI procedure will be used for the imputation.
1. data sets in total. In both cases, Hurley Stage at Baseline, Baseline antibiotic use, and value of the variable of interest lesion count at Baseline and at each post-Baseline visit (prior to the time point of interest in chronological order, see notes below about visits to include for different analysis sets) will be included in the imputation model. The post-Baseline values will need to be specified in chronological order in the imputation model so that SAS® PROC MI imputes variables from left to right (eg, the Week 2 value will be first imputed using regression based on the Baseline value, and then Week 4 value will be imputed using regression based on Baseline and Week 2

values, etc). Note that lesion count at earlier visits will also be used as predictors for the model of lesion count at later visits. The resulting data sets for each treatment arm will be combined into one complete data set based on each of the 100 imputations.

Note: The imputation model **based on the MCMC method** will only allow **joint multivariate normal numeric variables**. Therefore, Hurley Stage at Baseline and Baseline antibiotic use will be re-coded as indicator variables. For Baseline antibiotic use, this will simply be 0 for Baseline antibiotic non-users and 1 for Baseline antibiotic users. For Hurley Stage at Baseline, this will be 0 for Hurley Stage II participants and 1 for Hurley Stage III participants. In order to achieve model convergence, Baseline antibiotic use may be dropped from the model. If convergence is still not obtained, then Hurley Stage at Baseline may also be dropped from the model. **Additionally, if a variable is dropped in order to allow convergence for one model in a study, that variable does not have to be dropped from other models in the study if the model converges without dropping the variable. In other words, model convergence should be evaluated for each efficacy variable independently.**

Note: The imputation of each lesion type (inflammatory nodule, abscess, draining tunnel, etc) will be performed separately. The 100 data sets obtained for each type will be merged by imputation number and subject number.

- For each complete imputed data set, the dichotomous responder variable (eg, HiSCR 0 or 1) ~~based on the imputed %ΔAN and draining tunnel (fistula/sinus tract) count~~ will be computed. Each complete imputed data set will then be analyzed based on the logistic regression model.

Note: For derivation of HiSCR response, the AN, inflammatory nodule, abscess, and draining tunnel (fistula/sinus tract) counts at Week 16 in the imputed data sets will be compared directly to the observed Baseline counts to determine response. If values outside of the pre-defined range of values for lesion count (<0) are imputed, they will be cut off as appropriate after the multiple imputation procedure but before deriving the responder variable. For example, an imputed draining tunnel (fistula/sinus tract) count of **-1** would be changed to 0 before deriving the HiSCR responder variable. Additional ranges for values for secondary and other endpoints are defined in Table 4-2.

Note: Standard rounding rules will also be applied to the imputed values of endpoints that can only take integer values (eg, abscess count). For example, if a study participant has an abscess count imputed as 2.4, this imputed value would be rounded down to 2. This rounding step is performed after the multiple imputation but before deriving the responder variable.

Table 4-2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
Lesion count ^a	0	--	Yes
DLQI total score	0	30	Yes
hs-CRP	LLOQ/2	--	No
HSSDD item score	0	10	Yes

Table 4–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
HSSQ item score	0	10	Yes
HHS4	0	—	No
HiSQOL total score	0	68	Yes
HiSQOL symptom status score	0	16	Yes
HiSQOL psychosocial impact score	0	20	Yes
HiSQOL impact on physical activities score	0	32	Yes
EQ-5D-3L VAS	0	100	Yes
WPAI dimension scores	0	100	No for variables: “Percent work time missed due to problem” and “Percent overall work impairment due to problem”. Yes for variables: “Percent impairment while working due to problem” and “Percent activity impairment due to problem”. These two variables can only take values that are multiples of 10.
PHQ-9	0	27	Yes

^a Lesion counts will be imputed separately for each lesion type (abscesses, draining tunnels [fistulas/sinus tracts], inflammatory nodules, non-draining tunnels [fistulas/sinus tracts], non-inflammatory nodules, HS scars). **The imputed lesion counts will be used to derive the endpoints that are dependent on the lesion count data (eg, HiSCR₅₀).**

7. **Estimates of the adjusted responder rate for each treatment group and the associated SE are obtained** The Week 16 results **from the specified statistical analysis** (logistic regression model per Section 8.2.2) of each of the 100 imputed data sets. **These estimates will be combined for overall inference using Rubin’s rules, which account for the uncertainty associated with the imputed values (Rubin, 1987), and the combined estimates and SEs will be used to construct 95% CIs using the logit scale. This will be done using SAS PROC MIANALYZE. The combined estimates and 95% CIs on the logit scale will be back-transformed using the inverse logit link function to obtain the adjusted responder rates (%) and associated 95% CIs. This will be done using SAS PROC MIANALYZE.**

Note: The (unadjusted) proportion of responders will be calculated at each time point by treatment group from the imputed datasets using SAS PROC FREQ. These results will also be combined into an overall inference using SAS PROC MIANALYZE.

Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. However, in the event that there are computational challenges with the imputation model (eg, due to a standard deviation of

0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Some key points to consider relative to the calculation of the odds ratios and corresponding confidence intervals are noted below:

As the estimates of the odds ratios from the logistic regression models in Step 3 follow a log-normal distribution, a log transformation is needed to normalize these 100 odds ratio estimates. That is because the procedures for combining results from multiple imputed datasets assume that the statistics estimated from each imputed dataset are normally distributed. Therefore, the log of the odds ratio estimates from the logistic regression model are used when combining into a single inference (**Step 3**). **Additionally, the SE for the odds ratios are transformed as follows:** the use of PROC MIANALYZE in step 3). Appropriate transformations to the standard errors and p-values will also be made in order to get the correct confidence intervals. For the logistic regression using the p-value for the general association the Wilson-Hilferty transformation will be used (Ratitch, 2013).

$$SE = \frac{\log(UCL) - \log(LCL)}{2Z_{\alpha/2}}$$

Where UCL and LCL are the upper and lower confidence limit, respectively, for the CI of the odds ratio from the logistic regression model, and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). The estimates of the log odds ratio for Bimekizumab relative to placebo and the corresponding upper and lower CLs will be provided. The odds ratio will be then estimated by exponentiating the estimate of the log odds ratio. The odds ratio and the confidence limits of the odds ratio will be estimated as follows:

$$OR = \exp(\text{Log odds ratio estimate})$$

$$LCL = OR * \exp(-SE * Z_{\alpha/2})$$

$$UCL = OR * \exp(SE * Z_{\alpha/2})$$

Where OR is the back-transformed estimate of the odds ratio just described, SE is the SE of the log odds ratio and $Z_{\alpha/2}$ is the relevant critical value from the standard normal distribution (2.24 for a 97.5% CI). These calculations will be done such that odds ratios and corresponding CIs are calculated for the odds ratio of bimekizumab vs. placebo.

Note: . If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose, with a corresponding $Z_{\alpha/2}$ of 1.96.

In addition to calculating the odds ratio, associated CIs, and p-values for the pairwise comparisons of bimekizumab and placebo, the estimated proportion of responders (ie, estimated responder rate) and the difference in the proportion of responders between each bimekizumab treatment group and placebo will be estimated, and 2-sided 95% CIs will be

created for each difference. The creation of the estimates of the differences will be completed for each bimekizumab treatment group using the process detailed below:

8. Use the logistic regression model to calculate:

Least squares mean estimates of the log odds of bimekizumab (G_B) and placebo (G_P), as well as their corresponding standard errors (S_B and S_P , respectively).

Standard error of the least squares mean estimate of the log odds ratio (S_R)

9. Compute estimates for predicted proportions using the following transformations:

$$P_B = \exp(G_B) / (1 + \exp(G_B))$$

$$P_P = \exp(G_P) / (1 + \exp(G_P))$$

The difference in proportions is then given by:

$$D = P_B - P_P$$

10. Estimate the standard error of D by:

$$S_D = \sqrt{P_B^2(1-P_B)^2S_B^2 + P_P^2(1-P_P)^2S_P^2 + P_B(1-P_B)P_P(1-P_P)S_R^2 - P_B(1-P_B)P_P(1-P_P)(S_B^2 + S_P^2)}$$

~~The MCMC method for multiple imputation, as previously outlined, Missing data for continuous components of the primary endpoint and binary secondary efficacy endpoints will be imputed using MI as appropriate.~~

~~The above describes the procedure for binary endpoints. For continuous endpoints, the MI procedure will be similar to that described above with the following differences:~~

- ~~11. The absolute value of the given variable will be imputed. Once imputation has been performed across the 100 iterations specified, any values outside of the range of the given variable will be truncated accordingly.~~
- ~~12. The change from Baseline values will be computed based on the complete data sets.~~
- ~~13. The analysis model will be based on ANCOVA (Section 8.3.3 and Section 8.3.4) as opposed to logistic regression.~~

~~For other efficacy variables, MI will be used to account for missing values. The calculation steps impute missing data when possible and where specified. If the imputation model cannot converge, LOCF will be used. The MI procedure will also be similar to that described above will be based on the results provided from the logistic regression model of the multiple imputed datasets. An estimate for the difference in proportions, D, and corresponding standard error, SD, will be computed for each of these datasets. The results from these analyses will be combined into a single estimate of the difference in predicted proportion of response and a 2-sided 95% CI interval using SAS PROC MIANALYZE.~~

~~Note that this procedure indicates that the imputation model will be applied for each treatment group separately in PROC MI, which will be the default method. , for continuous and binary endpoints respectively. However, in the event there are computational challenges with the imputation model (eg, due to a standard deviation of 0 for responses of a given imputation), it is acceptable to modify the imputation model to include treatment as a variable in the model rather than running a separate model for each treatment group. It~~

should be noted that doing so assumes that treatment does not interact with any of the other variables in the imputation model.

Continuous endpoint

For continuous endpoints (eg, Change from Baseline in DLQI total score at Week 16), the MI method will be applied as follows:

5. The MCMC/monotone regression method described above in Step 1 for binary endpoints will be performed.
6. Based on the multiply imputed data sets obtained for the given variable, the change from Baseline will be derived for each of the 100 complete no inferential statistics will be calculated for the imputed data sets based on the observed Baseline value and the observed/imputed post-Baseline values. Note that if the value itself is being summarized, no additional derivation is needed.
7. If a statistical model is being used for the analysis of the variable, then that will be performed for each imputation in this step. If no statistical model is being used, then simple descriptive statistics will be calculated.
8. For data excluding hs-CRP, the following rules apply. The results of the 100 imputed data sets (based on the statistical model or descriptive statistics) are combined with means and standard errors Means and standard errors will be calculated using Rubin's rules (via PROC MIANALYZE). **Note that for the calculation of other descriptive statistics such as the median, min, and max, Rubin's rules do not apply. MI estimates will be computed by simply averaging the estimates from the multiple repetitions of the imputation algorithm. Thus, for median, Q1, Q3, minimum and maximum, the following approach will apply:** that will be used when summarizing continuous secondary efficacy variables by subgroup.
 - The data will be summarized by treatment, visit, and imputation, and the summary statistics will be computed.
 - Results will be summarized by taking the mean value of each summary statistic at each visit across all imputations.
 - The number of decimal places will remain the same as the original for display purposes (ie, if the mean was presented to 1 decimal place, then the mean of the means will also be presented to 1 decimal place).

For hs-CRP only, the following rules apply. The hs-CRP data will be presented using the geometric mean, 95% CI for the geometric mean, median, Q1, Q3, minimum and maximum. The change from Baseline will be expressed as the ratio to Baseline in the summaries. The following approach will be applied:

- Following the MI procedure, the ratio to Baseline will be calculated for any of the imputed values
- The natural logarithm of the absolute values and of the ratios to Baseline will be calculated

- The logged values will be summarized (using PROC MEANS) by treatment, visit and imputation
- The datasets will be combined using PROC MIANALYZE in order to get the mean and 95% CI estimates from the absolute values and ratios to Baseline (based on logged data) across imputations
- The estimates of the mean and 95% CI will be back-transformed to obtain the geometric mean and 95% CI on the original scale
- For the median, Q1, Q3, minimum and maximum the procedure outlined above for the other endpoints will be followed

If the imputation model cannot converge, LOCF will be used.

Change #19

Section 4.2.2.2 MI – MCMC/ Referenced-based imputation

The steps of the procedure were updated to:

The steps for the procedure are as follows:

1. For non-monotone (intermittent) missing data, MCMC will be used to impute lesion count data, with Baseline antibiotic use, Hurley Stage at Baseline, and lesion count at Baseline and at each post-Baseline visit (in chronological order) being included in the imputation model. This will be done only once for each participant in order to provide a dataset with monotone missing data.
2. Data will be processed sequentially by repeatedly calling SAS® PROC MI to impute missing outcome data at visits $t=1, \dots, T$, where T is Week 16 for HiSCR₅₀.
 - a. *Initialization.* Set $t=1$ (Baseline visit)
 - b. *Iteration.* Set $t=t+1$. Create a data set combining records from bimekizumab- and placebo-treated participants with columns for covariates (Hurley Stage at Baseline and Baseline antibiotic use) and outcomes at visits 1 to t . Outcomes for all bimekizumab-treated participants are set to missing at visit t and set to observed or previously imputed values at visits 1 to $t-1$. Outcomes for placebo-treated participants are set to observed at visit t or observed or previously imputed values at visits 1 to $t-1$. **The outcomes should be sorted in chronological order in the model.**
 - c. *Imputation.* ~~Run MCMC to impute~~ **Impute missing values** for visit t using previous outcomes for visits 1 to $t-1$, Baseline antibiotic use, and Hurley Stage at Baseline. Note that only placebo data will be used to estimate the imputation model since no outcome is available for bimekizumab-treated **participants at visit t . Consequently, the input dataset should include all study participants from placebo but only study participants from the bimekizumab arm that have values at timepoint t missing.**
 - d. Repeat steps 2a-2c, 100 times with different seed values (seeds ranging from 853 to 952) to create 100 imputed complete data sets. **Study participants whose missing values were imputed in the last PROC MI call will be included in the input dataset for the next PROC MI call. Standard rounding rules will be applied to the imputed values. If the MI procedure yields a value outside of**

the pre-defined range for the given variable, the value will be updated to be within the predefined range of values for the endpoint of interest. For example, the imputed value for inflammatory nodules will be updated to 0 in the case of an imputed value less than 0.

- e. *Analysis.* For each completed data set, the outcome (response or change from Baseline) will be calculated using the complete datasets (with no missing data).
3. Each complete imputed data set will then be analyzed based on the statistical model specified in this study (logistic regression). The Week 16 results from logistic regression of each of the 100 imputed data sets will be combined for overall inference using Rubin's rules, which account for the uncertainty associated with the imputed values (Rubin, 1987). This will be done using SAS PROC MIANALYZE.

Change #20

Section 4.2.2.3 Tipping Point Analysis

The steps for performing the tipping point analysis were updated:

Tipping point analyses will also be implemented as a supportive analysis for the primary efficacy endpoint.

The objective of the tipping point analyses is to **evaluate the sensitivity of results to departures from the missing at random assumption and to identify the point at which departures cause results to "tip" from statistically significant to statistically non-significant.** ~~As such, these identify assumptions about the missing data under which the conclusions from the main analysis change, ie, under which there is no longer evidence of a treatment effect. These tipping point analyses will only be performed if the primary efficacy analysis results in a statistically significant treatment effect ($p < 0.025$). Note that each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo independently for these analyses. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the p-value in this analysis method will be 0.05 instead of 0.025 throughout for that dose.~~

For tipping point analyses, data for participants after As a first step, the intercurrent event date (See Section 3.9) will be changed to missing prior to imputation but will not be changed to non-response after imputation.

The worst-case scenario will be evaluated first. All missing primary endpoint values for (HiSCR₅₀ at Week 16). Specifically, all study participants with a missing HiSCR₅₀ at Week 16 who have been randomized to bimekizumab (where missing values include observations after the intercurrent event date and any other missing values) will be imputed as non-responders, while all missing values for placebo-randomized study participants with a missing HiSCR₅₀ at Week 16 will be imputed as responders. While there is little justification for such an approach, it makes the most putative assumption possible against a bimekizumab treatment effect. After applying this imputation approach, a logistic regression model consistent with the one described for the primary analysis will be applied. If the p-value for the odds ratio of bimekizumab versus placebo remains significant is less than 0.025 for the particular bimekizumab dose regimen, then no further tipping point analyses are needed.

If this analysis based on the worst-case scenario results in a p-value **that is not significant** (eg, greater than 0.025), then additional tipping point analyses will be **performed to identify**. ~~Several assumptions will be made about average outcomes among the point at which results switch or~~ **“tip” from significant to non-significant. Note that subsets of study participants who prematurely discontinued study treatment and hence have a monotone missing data pattern (O’Kelly, 2014). In practice, it implies different delta adjustments will be made to the assumed responses on the monotone missing data in each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo independently for these analyses. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the with various degrees of plausibility with the goal to find for each treatment group the “tipping point” that will significantly reverse the primary result that yielded a p-value in this analysis method will be 0.05 instead of 0.025 throughout for that dose. In the tipping point analysis, a shift parameter or delta adjustment is applied to missing, and subsequently imported primary endpoint values (where missing values include observations after the intercurrent event and any other missing values). These delta adjustments will be done on the lesion count and will be implemented on the primary endpoint as follows:**

- 10. Data after intercurrent event date (See Section 3.9) will be set to missing.**
- 11. The same MCMC method described in Section 4.2.2.1 (Step 1a) will be implemented for non-monotone (intermittent) missing pattern values, using the same imputation model. This will be based on 100 sets of imputations. This will be done only once for each study participant to provide a dataset with monotone missing data.**
- 12. Based on the 100 datasets obtained in Step 2, a monotone regression model will be applied (using the same imputation model as in Step 2) as described in Section 4.2.2.1 (Step 1b).1) while adjusting the imputed values by various delta adjustments. This will be based on 1 imputation.**
- 13. Delta adjustments will be made to imputed lesion count values at Week 16, independently in each treatment group as described below.**
- 14. Delta adjusted imputed values will be truncated so that they are within the range of allowable values for each component.**
- 15. Following the delta adjustments for the individual components lesion counts, of the composite endpoint HiSCR₅₀ will then be derived based on the delta-adjusted multiply imputed data sets obtained for each component endpoint of interest.**

~~Several scenarios will be considered to define these shift parameters. Once defined, the same shift parameter value will be applied on the imputed endpoint value for all visits. Scenario 1 will assume that study participants randomized to bimekizumab and who have missing data have a lower probability of response compared to study participants randomized to placebo with missing data.~~

- ~~— For endpoints for which high scores are associated with a more favorable outcome, it will mean that:~~
 - ~~▪ A negative shift is applied to the imputed value for study participants randomized to bimekizumab to decrease the imputed value.~~

- ~~A positive shift is applied to the imputed value for study participants randomized to placebo to increase the imputed value.~~
- ~~For endpoints for which high scores are associated with a less favorable outcome, it will mean that:~~
 - ~~A positive shift is applied to the imputed value for study participants randomized to bimekizumab to decrease the imputed value.~~
 - ~~A negative shift is applied to the imputed value for study participants randomized to placebo to increase the imputed value.~~

~~For each continuous variable, a set of possible values will be first pre-defined for the shift parameter (example: 0, 1, 2, 3).~~

- 16. Each of the 100 imputed datasets will then be analyzed using a logistic regression model with factors of treatment group, , Baseline Hurley Stage, and Baseline antibiotic use.**
- 17. The results obtained from the 100 logistic regression analyses in Step 7 will be combined for overall inference using Rubin's rules, and the results obtained for each shift parameter will be presented in a single table.**
- 18. Steps 4 to 8 will be repeated so that, at each iteration, missing values are adjusted with a larger delta than at the previous iteration. Depending on the results obtained, shift parameters with more granularity (eg, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9) may also be investigated. The process will go on until the p-value for the odds ratio between bimekizumab and placebo is no longer statistically significant (eg, ≥ 0.025). The odds ratio, 97.5% CI (or 95% depending on the significance level being used for testing), and p-values obtained for each value of delta will be combined in one single table.**

The delta adjustments result in study participants randomized to bimekizumab with missing data having a lower probability of response compared to study participants randomized to placebo with missing data. Since HiSCR₅₀ response is an endpoint for which high lesion counts are associated with a less favorable outcome:

- **A positive adjustment is applied to the imputed value for study participants randomized to bimekizumab in order to increase the imputed value and decrease the likelihood of response.**
- **A negative adjustment is applied to the imputed value for study participants randomized to placebo in order to decrease the imputed value and increase the likelihood of response.**

To start, imputed values within each values within each lesion type, will be adjusted by the same value in each treatment arm. This adjustment will be 5% of the observed range within that lesion type. Depending on the results obtained, this adjustment will be multiplied for step 9 above (2 times, 3 times the initial adjustment) until the p-value is no longer statistically significant. This can be an adjustment of preselected integer values (eg, 1, 2, and 3) or adjustments at intervals equal to a percentage of the allowable range of the component (eg, 5% of range of 10 to give 0.5, 1, 1.5 etc.). Depending on the results obtained, more granular adjustments (eg, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9) may also be implemented to better

~~identify the point at which results "tip". More robust primary analysis results will require larger adjustments to tip the results from significant to insignificant.~~

Additionally, study participants randomized to bimekizumab with an intercurrent event should be set to non-response, after applying the delta adjustment outlined in Step 6 above. This ensures study participants randomized to bimekizumab do not have a higher probability of response in the tipping point analyses compared to the primary analysis (ie, a study participant randomized to bimekizumab who is non-responder in the primary analysis cannot become a responder in the tipping point analyses).

Change #21

New section was added.

Rationale for estimand (Section 4.2.3):

Intercurrent events have been identified within the estimands for this study because their potential to impact efficacy assessments linked with the primary and secondary study objectives. In order to account for the effect of any observed post-randomization intercurrent events on the efficacy analyses, the following estimand strategies will be implemented when evaluating the primary and secondary efficacy endpoints:

- A composite estimand strategy will be used for the primary analysis of the binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, flare),
- A hypothetical estimand will be used for the primary analysis of the continuous secondary endpoints (change from Baseline in DLQI total score and in "worst pain" item for the HSSDD).

Change #22

New section was added.

Composite estimand (Section 4.2.3.1):

A composite estimand strategy as defined in Section 8.2.2 allows incorporation of the two intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) within the definition of the endpoint. These intercurrent events are considered meaningful to the efficacy outcome following receipt of study medication. For example, within the proposed composite estimand framework, a randomized study participant who discontinues from study treatment due to lack of efficacy prior to Week 16 will be considered a treatment failure at Week 16 regardless of the lesion count assessment performed at that visit.

The assumptions and robustness of the primary analysis (modified composite estimand as defined in Section 8.2.2) will be assessed through the sensitivity analyses defined in Section 8.2.3. The impact of intercurrent event handling and data imputation methods on endpoint derivation will also be assessed via the analyses of lesion counts and derived HiSCR variables as specified in Section 8.4.2.1 and Section 8.4.1.1, respectively.

Change #23

New section was added.

Hypothetical estimand (Section 4.2.3.2):

The hypothetical estimand is defined in Section 8.3 and involves a data-driven approach to account for the potential impact of intercurrent events (eg, receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy) on the analysis of continuous efficacy endpoints. Under this framework, outcomes for study participants without an intercurrent event are analyzed as observed. Conversely, outcomes for study participants with an intercurrent event are imputed via a multiple imputation model, ie any recorded data on or after the intercurrent event will be set to missing and imputed via multiple imputation following the strategy established in Section 4.2.2.1.

Change #24

Section 4.2.4 Dates and times

Partial stop and end date imputation rules were updated:

- Imputation of Partial Start Dates
 - If only the month and year are specified:
 - **If the month and year of first dose of study medication is the same as the month and year of the partial start date, then use the date of first dose of study medication,**
 - **Else, if the month and year of the partial start date are the same as the month and year of a study medication switch date, then use the date of study medication switch,**
 - **Otherwise, use the 1st of the month of the partial start date;**
 - If only the year is specified:
 - **If the year of first dose of study medication is the same as the year of the partial start date, then use the date of first dose of study medication,**
 - **Else, if the year of the partial date is the same as the year of a study medication switch date, then use the date of study medication switch,**
 - **Otherwise, use the 1st of January of the year of the partial start date;**
 - If the start date is completely unknown:
 - **If the stop date is unknown or not prior to the date of first dose of study medication, then use the date of first dose of study medication,**
 - **If the stop date is prior to the date of first dose of study medication, then use the 1st of January of the year of the stop date.**
- Imputation of Partial Stop Dates
 - **If only the month and year are specified, :**
 - **Use the last day of the month of the partial stop date;**
 - **If only the year is specified**
 - **use December 31st of the year of the partial stop date;**

- If the stop date is completely unknown,
 - Do not impute the stop date.

Note that if the stop date or the imputed stop date is prior to the imputed start date, then follow the procedure outlined below:

- If only the year of the start date is specified:
 - If the year of start date is the same as the year of first dose of study medication and the imputed stop date is after the date of first dose of study medication, then set the start date to the date of first dose of study medication,
 - Otherwise, set the 1st January of the year of the start date;
- If only the month and year of start date are specified:
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is on or after the date of first dose of study medication then set the start date to the date of first dose of study medication,
 - If the month and year of the start date is the same as the month and year of first dose of study medication and the imputed stop date is before the date of first dose of study medication then set the start date to the 1st of the month of partial start date.

Change #25

Section 4.6 Use of an efficacy subset of participants

The section was updated to:

A sensitivity analysis of the primary endpoint will be performed based on the FAS, the PPS, and the CFS.

Change #26

Section 4.8 Examination of subgroups

This section was updated to:

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, and flare endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period separately. **Additional subgroup analyses will be performed on the change from Baseline in the worst pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.**

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, and flare which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

The following **subgroup variables** will be determined using Baseline data, **except for analgesic use, lesion intervention, and antibody positivity**:

- Age (<40 years, 40 to <65 years, ≥65 years)
- Gender (male, female)

- Disease duration (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for analysis.

- Region (North America [Canada, USA], Western Europe [France, Germany, Ireland, Italy, Spain, United Kingdom], Central/Eastern Europe [Bulgaria, Czech Republic, Hungary, Poland], Asia/Australia [Australia, Israel, Japan])
- Weight (≤100 kg, >100 kg)
- BMI (<25 kg/m², 25 to <30 kg/m², ≥30 kg/m²)
- Race (Black or African American, White, All Other Races [**American Indian or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other/Mixed**])
- Systemic antibiotic therapy at randomization (yes, no)
- Prior biologic therapy for any indication (yes, no)
- Prior biologic therapy for HS (yes, no)
- Hurley Stage at Baseline (II or III)
- Analgesic users (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period (Section 6.4.2 specifies how participants are classified as analgesic users)
- Lesion intervention (yes, no), separately for the Initial Treatment Period and the Maintenance Treatment Period
- Antibody positivity (confirmatory assay: negative or positive)

~~Any analgesic rescue medication taken during the study, lesion intervention (including new post-Baseline antibiotic use or dose adjustments) and antibody positivity are the only subgroups that are not determined by Baseline data. They will be presented in a separate table.~~

Subgroup analyses will also be performed by visit **The following subgroups for analysis on the change from Baseline in the worst pain score as measured by HSSDD and in the DLQI total score through Week 16.** ~~The following subgroups for analysis will be determined based on medication use during the Initial Treatment Period:~~

- Antihistamines users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users)
- Analgesics users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as analgesic users)
- Systemic antibiotic therapy start/increase after randomization during the Initial Treatment Period (yes, no)

All summaries will be based on imputed data as appropriate and will include descriptive statistics only.

Change #27

Section 5. 1 Study participant disposition

The following sentence was added:

Participants are defined as completing the Initial Treatment Period if they have a Week 16 visit, or if they fail to attend the Week 16 visit but attend at least one visit in the Maintenance Treatment Period.

The following summaries were also added:

To assess participant disposition (entry and periods in the study) during the COVID-19 pandemic, study participant disposition will also be assessed by period of the COVID-19 pandemic (pre – during – post), by comparing the dates of visits (or events) to the dates of the COVID-19 pandemic period. The dates to categorize the periods of the COVID-19 pandemic (pre/during/post) are defined below:

- Pre-COVID-19 pandemic period: Period prior to COVID-19 pandemic start date defined as 11-Mar-2020
- COVID-19 pandemic period: Period from 11-Mar-2020 through the COVID-19 pandemic end date which is currently not defined at the time of approval of the SAP
- Post-COVID-19 pandemic period: Period after the declaration of the end of the pandemic

Change #28

The following new section was added:

Impact of COVID-19 (Section 5.2)

A listing of visits affected by COVID-19 will be presented based on the ES including the visit, date of visit, relationship to COVID-19, impact category and a narrative (short description) of the event. These data will be summarized for non-randomized participants and by treatment group and overall, for enrolled participants.

A summary of study visits by COVID-19 pandemic period (pre – during – post) will be presented for participants enrolled prior to and during the pandemic.

In addition, in order to assess the potential impact of COVID-19 on the collection and reporting of efficacy data, a separate summary on the RS will be presented to display missing data as well as data collected via an alternative modality (e.g.: phone, video call) for efficacy endpoints included in the hierarchy (Section 4.5). For these displays, missing data will be presented only for visits affected by COVID-19, as reported on the dedicated eCRF page. Missing data at other visits and for other reasons will not be included. Note that the remote contingencies for COVID-19 or other exceptional circumstances are not applicable to efficacy assessments and documentation (eg, lesion-based assessments, photography) that require direct face-to-face physician/participant interaction.

Change #29

Section 5.3 Protocol deviations

The following text has been added:

A separate summary of participants with protocol deviations related to COVID-19 will be provided.

A by-participant listing of protocol deviations will be provided. Protocol deviations (eg, missing assessments or visits) related to COVID-19 will be listed separately.

Change #30

Section 6.2 Other Baseline characteristics

The last 5 bullets were updated:

- Duration of disease (<median, ≥median)

The median disease duration will be calculated based on all participants in the analysis set used for the summary.

- Baseline antibiotic use (yes, no) (**According to the randomization strata**)
- **Baseline antibiotic use (yes, no) (Derived)**
- Hurley Stage at Baseline (**According to the randomization strata**)
- **Hurley Stage at Baseline (Derived)**

The following text was added or the additional summaries:

In addition, the following Baseline disease characteristics will be summarized by **the derived** Baseline Hurley Stage and by **the derived** Baseline antibiotic use and treatment group for the RS

Change #31

Section 6.4 Prior and concomitant medications

The following sentence was updated:

Prior medications include any medications that started ~~before~~ prior to the start date of study medication. Concomitant medications are **any medication that has a start date on or after the start date of study medication, or any medication that has a start date on or before the last dose of study medication + 28 days (whether placebo or bimekizumab).** ~~medications taken at least 1 day in common with dosing period.~~

The following sentence was added:

Additional summaries for the Initial Treatment Period and Maintenance Treatment Period will be presented for participants taking systemic antibiotic medications that qualify as intercurrent events as described in Section 3.9.

Change #32

Section 8 Efficacy Analyses

This section was updated:

All efficacy analyses of primary, ~~and~~ secondary, **and other** variables will be performed on the RS unless otherwise specified. ~~All efficacy analyses of other efficacy variables will be performed on the RS and MS unless otherwise specified.~~ All efficacy summary tables will be displayed by treatment **sequence** unless otherwise specified. The primary and secondary endpoints, and their components, will also be summarized by **the derived** Hurley Stage at

Baseline (grouping each stage and overall) and treatment **sequence** and by **the derived** Baseline antibiotic use (yes/no and overall) and treatment group.

Change #33

Section 8.2 Primary efficacy endpoint

The following rows in Table 8-2 were added or amended:

Table 8-1: Estimand Details and Attributes for Primary Endpoint

		Estimands for Primary Endpoint			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Primary Objective: To evaluate the efficacy of bimekizumab in study participants with moderate to severe HS					
HiSCR ₅₀	Sensitivity (Section 8.2.3.1)	HiSCR ₅₀ response at Week 16	RS	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handles by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.

Table 8-1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.2)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a hypothetical strategy, whereby all data at and after the intercurrent event will be treated as missing. Composite strategy, as for the primary analysis. Composite strategy, as for the primary analysis.	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – Reference-Based Regression under a missing not at random assumption.
HiSCR ₅₀	Sensitivity (Section 8.2.3.3)	HiSCR ₅₀ response at Week 16	RS	Composite strategy ^a , as for the primary analysis.	A tipping point analysis will be used where various delta adjustments will be made to the assumed responses on the monotone missing data in each treatment group independently with various degrees of plausibility. The odds ratio versus placebo is based on a logistic regression for each value of delta.

Table 8-1: Estimand Details and Attributes for Primary Endpoint

Objective Clinical Category	Statistical Category (Section)	Estimands for Primary Endpoint			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HiSCR ₅₀	Sensitivity (Section 8.2.3.8)	HiSCR ₅₀ response at Week 16	CFS	Composite strategy, as for the primary analysis.	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.9)	HiSCR ₅₀ response at Week 16	RS	The same two intercurrent events used for the primary analysis will be used. Any missing data due to COVID-19 will also be considered an intercurrent event. A composite strategy will be used, ie, the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as treatment failures (nonresponders).	The odds ratio versus placebo is based on a logistic regression, as for the primary analysis.
HiSCR ₅₀	Sensitivity (Section 8.2.3.10)	HiSCR ₅₀ response at Week 16	RS	Composite strategy, as for the primary analysis.	The odds ratio versus placebo based on a stratified Cochran-Mantel-Haenszel (CMH) test. Missing values not preceded by an intercurrent event will be imputed using MI – MCMC/Monotone Regression under a missing at random assumption.

CFS=Covid-19 Free Set; CMH=Cochran-Mantel-Haenszel; COVID-19=coronavirus disease 2019; FAS=Full Analysis Set; HiSCR=Hidradenitis Suppurativa Clinical Response; IES=intercurrent event(s) strategy; MCMC=Markov Chain Monte Carlo; MI= multiple imputation; PLS=Population-level summary; Pop=Population; PPS=Per-Protocol Set; RS=Randomized Set

^a The composite estimand strategy will be modified in the tipping point analysis such that participants with intercurrent events will be treated as nonresponders only in the bimekizumab treatment groups.

Change #34

Section 8.2.2 Primary analysis of the primary efficacy endpoint

The following text was updated:

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, and Baseline antibiotic use. The odds ratio versus placebo, p-value (from Wald test), and **97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96 .**

Change #35

New section was added:

Analysis on CFS (Section 8.2.3.8)

The primary efficacy analyses from Section 8.2.2 will be repeated based on the CFS.

Change #36

New section was added:

Analysis including COVID-19 as intercurrent event (Section 8.2.3.9)

An additional sensitivity analysis will include an additional intercurrent event. The composite estimand combines the clinically meaningful improvement from Baseline based on the HiSCR₅₀ response and completion of study treatment through Week 16 without receiving systemic antibiotic rescue medication or discontinuing IMP due to an AE or lack of efficacy.

The following 4 attributes describe the estimand that will be used to define the treatment effect of interest for this sensitivity efficacy analysis:

5. Population=Study participants meeting the protocol-specified inclusion/exclusion criteria.
6. Study participant-level outcome=HiSCR₅₀ at Week 16.
7. Intercurrent event handling=An intercurrent event is defined as receipt of systemic antibiotic rescue medication, discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16, or missing data due to COVID-19. More information is provided in Section 3.9. A composite strategy will be implemented in which a positive clinical outcome is defined as achieving HiSCR₅₀ at Week 16 and not receiving systemic antibiotic rescue medication, not discontinuing study treatment due to an AE or lack of efficacy through Week 16, and not having missing data due to COVID-19. Any missing data at Week 16 that are not preceded by an intercurrent event will be imputed using multiple imputation as defined in Section 4.2.1.
8. Population-level summary measure=conditional odds ratio comparing bimekizumab to placebo adjusted for stratification factors used in the randomization.

The same imputation techniques and analysis model as in the primary efficacy analyses will then be used.

Change #37

Section 8.3 Secondary efficacy endpoints

The following text has been added:

Sensitivity analyses of the secondary endpoints will be performed on the CFS.

Change #38

Section 8.3.2 Flare by Week 16

The following text was deleted:

Disease flare status will be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in each treatment group. A bar chart of percentage of subjects with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Initial Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Initial Treatment Period will be presented.

Change #39

Section 8.3.3 DLQI Total Score at Week 16

The following paragraph was updated to:

Change from Baseline in DLQI total score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use and Baseline value as a covariate. The least square mean (LSM), standard error (SE), **and 95% CI for the LSM** will be presented by treatment group. For the comparison between placebo and bimekizumab: the difference between the LSM, the associated **97.5% CI** for the contrasts, and the corresponding p-value **will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96.**

Change #40

Section 8.3.4 Skin Pain score at Week 16, as assessed by the “worst pain” item in the HSSDD

This section was updated:

The items on the HSSDD assess patients' perception of the core symptoms of HS experienced in the past 24 hours: pain, smell or odor, drainage or oozing from HS lesions, and itch on an 11-point numeric rating scale (NRS). Two items assess skin pain: worst skin pain and average skin pain.

Weekly averages will be derived for each of the items of the HSSDD **for weeks matching the post-Baseline dosing weeks** up to Week 16. A weekly average is the sum of the scored item over the course of the study week divided by the number of days on which the item was completed. The weekly averages will be relative to the respective visit date except for Baseline, which will be anchored to the first dose of study drug. A weekly average will only be calculated if at least 4 non-missing values (not necessarily consecutive) are available. Otherwise, the HSSDD weekly average for the given question will be set to missing.

Baseline will be computed as the average from the first 7 consecutive day period in which there are at least 4 non-missing entries. That is, first consider the first 7 consecutive days prior to the Baseline visit, but not including the Baseline visit day itself. If there are at least 4 non-missing values (not necessarily consecutive), then the Baseline average will be calculated. If there are less than 4 values, the 7 consecutive day period will move one day earlier. If there are at least 4 non-missing values (not necessarily consecutive) in that period, then the Baseline average will be calculated. This will continue until there are at least 4 non-missing values in a 7 consecutive day period in the 14 days prior to Baseline. If there is no period in which there are at least 4 non-missing entries, then the Baseline value will be set to missing. ~~Baseline will be computed as the average from the 2 weeks prior to Baseline, up to and including the data from the Baseline visit. If less than 7 non-missing values are available for a given question, the Baseline for the given question will be set to missing.~~

Change from Baseline in worst skin pain score is defined as the average Week 16 worst skin pain score minus the Baseline worst skin pain score. Missing data imputation described in Section 4.2.1.2 will be applied to the weekly averages and not to the individual daily PRO data.

Change from Baseline in worst skin pain score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use, analgesic use (Section 6.4.2) and Baseline value as a covariate. A treatment-by-analgesic-use interaction term will also be added to the model and removed if not significant.

The LSM, SE, and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab, the difference between the LSM, the associated 97.5 95% CI for the contrasts, and the corresponding p-value will be presented. **If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose** with a corresponding $Z_{\alpha/2}$ of 1.96.

Change #41

Section 8.4.1.2 Time to response of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀

The following text was updated:

Initial Treatment Period

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first HiSCR_{xx} response, Date of Week 16 visit) – Date of **first dose of study medication** Baseline visit + 1, here xx represents 25, 50, 75, 90, 100 respectively. All visits including unscheduled visits are considered.

Participants who discontinue study treatment without achieving a given HiSCR response prior to Week 16 visit will be censored at the date of **last lesion count assessment**. ~~discontinuation.~~

Participants who reach the Week 16 Visit without achieving the given response will be censored at the date of the Week 16 Visit. Participants who experience an intercurrent event **prior to achieving a HiSCR response** will be censored at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

The following text was deleted:

Combined Initial and Maintenance Treatment Period

An additional time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the combined Initial and Maintenance Treatment Period will be calculated as above, where the Week 48 visit is considered instead of Week 16.

Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the combined Initial and Maintenance Treatment Period will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group. ~~These summaries will be limited to participants randomized to bimekizumab.~~

Kaplan-Meier plots of time to HiSCR responses will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 (since there are no responders at Week 0) and will increase over time, representing time to achieving the response.

The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.

Maintenance Treatment Period

~~For participants randomized to placebo, an additional time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) during the Maintenance Treatment Period will be calculated as:~~

~~Min (Date of first HiSCR_{xx} response, Date of Week 48 visit) — Date of Week 16 visit + 1, here xx represents 25, 50, 75, 90, 100 respectively. All visits including unscheduled visits are considered.~~

~~Participants who discontinue study treatment without achieving a given HiSCR response prior to Week 48 visit will be censored at the date of discontinuation. Participants who reach the Week 48 Visit without achieving the given response will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event will be censored at the date of the intercurrent event. Participants will be censored at Week 16 if there is no Post Week 16 lesion count assessment.~~

~~Time to HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response during the Maintenance Treatment Period will each be estimated and presented using the Kaplan-Meier product limit method for the placebo/bimekizumab 320mg Q2W treatment group.~~

~~The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.~~

Change #42

New section was added:

HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ among Week 16 Responders (Section 8.4.1.4)

See Section 8.4.1.1 for the derivation of HiSCR response.

Summaries of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at each visit from Week 16 through Week 48 will be summarized based on a subset of participants in the MS who achieve response at Week 16. The summaries will be as follows:

- HiSCR₅₀ responder rate based on participants who achieved HiSCR₅₀ response at Week 16
- HiSCR₇₅ responder rate based on participants who achieved HiSCR₇₅ response at Week 16
- HiSCR₉₀ responder rate based on participants who achieved HiSCR₉₀ response at Week 16
- HiSCR₁₀₀ responder rate based on participants who achieved HiSCR₁₀₀ response at Week 16

Line plots of the above HiSCR responder rate categories over time (from Week 16 to Week 48), by treatment group, will be produced.

Change #43

Section 8.4.1.5 Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ in Week 16 responders

The following text has been updated:

See Section 8.4.1.1 for the derivation of HiSCR response.

Time to loss of response will be based on the MS and include only participants who had the corresponding HiSCR response at Week 16 (considering intercurrent event handling from the composite estimand described in Section 8.2.2).

Time to loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ (in days) is defined as: Date of loss of response of HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ - Date of Week 16 treatment **administration** + 1.

Time to loss of response will be estimated and presented using the Kaplan-Meier product-limit method for each treatment. Participants who experience an intercurrent event **prior to loss of response** will be considered as having lost response on the date of intercurrent event.

Participants who reach the Week 48 Visit without loss of response will be censored at the date of the Week 48 Visit. Participants who discontinue treatment or study, for reasons other than those already defined for an intercurrent event, and who have not yet displayed loss of response by the time of withdrawal, will be censored at the date of **the last lesion count assessment withdrawal**.

~~The summary for each HiSCR response will include only participants who had the corresponding HiSCR response at Week 16.~~

Change #44

Section 8.4.1.6 Partial Response

The following paragraph was updated:

The number and percentage of participants who are partial responders at Week 16 and become HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders in the Maintenance Treatment Period will be summarized by treatment group and visit. These analyses will be based on the subset of participants in the **MS RS** that are partial responders but not HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders, respectively, at Week 16. **These summaries will be based on observed case data and will not consider the occurrence of intercurrent events.**

Change #45

New section was added:

Flare relative to Baseline (Section 8.4.3)

See Section 8.3.2 for the derivation of flare.

Disease flare status will be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in each treatment group. This summary will also include the number of participants with any flare in the Initial Period, Maintenance Period, and the combined Initial and Maintenance Period. A bar chart of percentage of participants with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Initial Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Initial Treatment Period will be presented.

Change #46

The following section was deleted:

Flare by Week 48 (Section 8.4.3)

See Section 8.2.1 for the derivation of AN count.

Disease flare by Week 48 is defined when at least a 25% increase in AN count with an absolute increase of ≥ 2 AN relative to Week 16 is observed by Week 48. A participant's disease flare status (yes/no) will be determined at each visit in the Maintenance Treatment Period using these criteria and will be listed with the other lesion count assessment data in the data listings.

The number of participants who experience at least 1 disease flare by Week 48 will be summarized by treatment group.

Disease flare status during the Maintenance Period will also be summarized by treatment group and visit using frequencies and percentages. The denominator for the percentage calculations will be the number of participants with non-missing data in the Maintenance Treatment Period in each treatment group. A bar chart of percentage of subjects with flare, by visit and treatment, will be presented.

In addition, for each participant, the number of flares during the Maintenance Treatment Period will be calculated and summarized by treatment group. A corresponding histogram summarizing the number of flares during the Maintenance Treatment Period will be presented.

Change #47

Section 8.4.4 Time to flare by Week 16

This section was updated as follows:

See Section 8.3.2 for the derivation of flare ~~by Week~~.

Time to flare (in days) during the Initial Treatment Period will each be calculated as:

Min (Date of first flare, Date of Week 16 visit) – Date of **first dose of study medication**
~~Baseline visit~~ + 1. All visits in the Initial Treatment Period including unscheduled visits are considered.

Participants who discontinue study treatment without experiencing a flare prior to Week 16 Visit will be censored at the date of **last lesion count assessment**. ~~discontinuation~~. Participants who reach the Week 16 Visit without experiencing a flare will be censored at the date of the ~~the~~ Week 16 Visit. Participants who experience an intercurrent event **prior to experiencing a flare** will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will be estimated and presented using the Kaplan-Meier product-limit method for each treatment group.

Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.

The median time to **flare response**, including the 2-sided 95% confidence interval, will be calculated for each treatment. Comparisons of bimekizumab vs placebo will be analyzed using a log-rank test stratified by Hurley Stage at Baseline and Baseline antibiotic use.

Change #48

Section 8.4.5 Time to flare by Week 48

This section was updated as follows:

See Section 8.3.2 for the derivation of flare **relative to Baseline** by Week 48.

Maintenance Treatment Period

~~Time to flare (in days) during the Maintenance Treatment Period will each be calculated as:~~

~~Min (Date of first flare, Date of Week 48 visit) – Date of Week 16 visit + 1. All visits in the Maintenance Treatment Period including unscheduled visits are considered.~~

~~Participants who discontinue study treatment without experiencing a flare prior to Week 48 visit will be censored at the date of discontinuation. Participants who reach the Week 48 Visit without experiencing a flare will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Week 16 if there is no Baseline lesion count assessment or no Post Week 16 lesion count assessment.~~

~~Time to flare will each be estimated and presented using the Kaplan-Meier product limit method for each treatment group.~~

~~Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.~~

~~The median time to response, including the 2-sided 95% confidence interval, will be calculated for each treatment.~~

Combined Initial and Maintenance Treatment Period

Time to flare (in days) during the combined Initial and Maintenance Treatment Period will be calculated as:

Min (Date of first flare, Date of Week 48 visit) – Date of **first dose of study medication** ~~Baseline visit~~ + 1. All visits in the up to Week 48 including unscheduled visits are considered.

Flare will be defined relative to the Baseline visit. Participants who discontinue study treatment without experiencing a flare prior to Week 48 visit will be censored at the date of **last lesion count assessment**. ~~discontinuation~~. Participants who reach the Week 48 Visit without experiencing a flare will be censored at the date of the Week 48 Visit. Participants who experience an intercurrent event **prior to experiencing a flare** will be treated as experiencing a flare at the date of the intercurrent event. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline lesion count assessment.

Time to flare will each be estimated and presented using the Kaplan-Meier product-limit method for each treatment group. ~~This summary will be limited to participants randomized to bimekizumab.~~

~~Kaplan-Meier plots of time to flare will be presented by treatment group. In these Kaplan-Meier plots, the line will start at 0 and will increase over time, representing time to achieving the response.~~

The median time to **flare response**, including the 2-sided 95% confidence interval, will be calculated for each treatment.

Change #49

Section 8.4.6 International Hidradenitis Suppurativa Severity Score System (IHS4)

The following sentence was added:

The observed IHS4 score, change and percentage change from Baseline will be summarized by treatment group and visit. **Missing IHS4 scores will be imputed using the multiple imputation procedure specified in Section 4.2.2.1, where IHS4 scores will be derived based on the imputed lesion counts.**

Change #50

Section 8.4.10 Time to initiation of systemic rescue therapy in the Initial Treatment Period

The following text was updated:

See Section 3.9 for the definition of a systemic antibiotic rescue therapy.

Time to initiation of systemic rescue therapy (in days) during the Initial Treatment Period will be calculated as:

Min (Date of initiation of rescue therapy, Date of change in the dose/type of current antibiotic, Date of Week 16 visit) – Date of **first dose of study medication** ~~Baseline visit~~ + 1.

Participants who discontinue ~~the study treatment~~ without initiating systemic rescue therapy prior to Week 16 visit will be censored at the date of discontinuation. Participants who reach the Week 16 Visit without initiating systemic rescue therapy will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no ~~Baseline lesion count assessment or no Post-Baseline visit lesion count assessment~~.

Change #51

Section 8.4.11 Time to an intercurrent event in the Initial Treatment Period

The following text was updated:

See Section 3.9 for the definition of an intercurrent event.

Time to an intercurrent event (in days) during the Initial Treatment Period will be calculated as:

Min (Date of **intercurrent event** initiation of rescue therapy, ~~Date of change in the dose/type of current antibiotic, Date of withdrawal due to AE or lack of efficacy, Date of Week 16 visit~~) – Date of **first dose of study medication** Baseline visit + 1.

Participants who discontinue the study treatment without experiencing an intercurrent event prior to Week 16 visit will be censored at the date of discontinuation. **That includes participants who discontinue from the study for reasons other than Adverse Event and Lack of Efficacy.** Participants who reach the Week 16 Visit without experiencing an intercurrent event will be censored at the date of the Week 16 Visit. Participants will be censored at Baseline if there is no Baseline lesion count assessment or no Post-Baseline visit lesion count assessment.

Change #51

Section 8.4.12 Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)

The following sentence was updated:

Change from Baseline in Worst Pain score and Worst Itch score at Week 16 will additionally be summarized **by visit and** by analgesic and antihistamine use status (Section 6.4.2), respectively.

Change #53

Section 8.4.13 Hidradenitis Suppurativa Symptom Questionnaire (HSSQ)

The following text was updated:

HSSQ response for pain **item score** is defined as at least a 30% reduction and at least a 1-unit reduction from Baseline in HSSQ among study participants with a score of ≥ 3 at Baseline.

The number and percentage of responders for pain **item score** will be summarized by treatment group and visit **based on the MS.**

The number and percentage of participants who were responders at any timepoint in the Maintenance Treatment Period will be summarized by treatment group for the skin pain score **based on the MS.**

Change from Baseline in pain score and itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

Change #54

Section 8.4.15 Hidradenitis Suppurativa Quality of Life (HiSQOL)

The following text was updated:

Summary statistics of the actual values and change from Baseline values will be used to summarize HiSQOL **domain and total scores** for each visit by treatment group. The table will

display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3. **The imputed HiSQOL total score will be derived based on the imputed subscales.**

The number and percentage of participants that complete the HiSQOL will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS (or MS, as appropriate). The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit.

A by-participant listing of the HiSQOL questionnaire, HiSQOL responses, **domain and total scores** and change from Baseline will be provided.

Change #55

Section 9.1 Pharmacokinetics

The following sentence was updated:

~~All However, all~~ PK concentrations **collected will be listed irrespective of the dosing or sampling occurring out of window.**

Change #56

Section 8.2.3.1 Derivation of palmoplantar IGA response

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-tiered assay approach: screening assay, confirmatory assay, and titration assay. Samples confirmed as positive within the confirmatory assay will be further evaluated **for the presence of in a neutralizing anti-assay to evaluate the potential of the ADAb to neutralize the activity of bimekizumab antibodies specific to IL-17AA, IL-17FF (IL-17A or IL-17F, or both) in vitro.** Samples will be taken at Baseline, then at study Weeks 4, 8, 12, 16, 20, 24, 36 and 48, and at PEOT and SFU timepoints.

~~The screening~~ Screening, confirmatory, and titer cut **point will be used to determine points of the status of anti-bimekizumab antibodies in respective assays will be determined by the test sample as Positive Screen (PS) or Negative Screen (NS). For bioanalytical laboratory based either on commercially available drug naïve samples presenting anti-bimekizumab antibody levels that are PS, a further confirmatory assay will be performed, and the result of which will be reported as either Positive Immunodepletion (PI) or Negative Immunodepletion (NI).**

ADAb status for each sample will be derived as follows:

- **Sample values that are either NS, or PS and NI and where the bimekizumab concentration is less than the validated ADAb assay drug tolerance limit will be defined as anti-bimekizumab antibody negative.**
- **Sample values that are either NS, or PS and NI and where the bimekizumab concentration exceeds the validated ADAb assay drug tolerance limit will be defined as inconclusive.**

- **Sample values that are PS and PI will be defined as ADAb positive (regardless of availability of a titer value)**
- **Missing or non-evaluable samples will be defined as missing**

Positive immunodepletion samples will be titrated, and the ADAb titer (reciprocal dilution factor including minimum required dilution) will be reported. Subsequently, PI samples will also be subject to a neutralizing assay to evaluate the potential of ADAb to neutralize the target binding of bimekizumab (IL-17AA or IL-17FFIL17F or both) in vitro.

The following definitions will be applied regarding ADAb status of each test samples:

- ~~An ADAb status will be confirmed as positive for any sample with an ADAb level that is positive screen and positive immunodepletion.~~
- ~~An ADAb status of negative will be concluded for any sample with an ADAb level that is either negative screen or (positive screen and negative immunodepletion).~~

~~If the titer for an ADAb level that is positive screen and positive immunodepletion is missing, then a conservative approach will be used and ADAb status will be consider as positive. No imputation rules apply for the missing titer. If the ADAb level is positive screen but no confirmatory result could be determined, then a conservative approach will be used and ADAb status will be consider as positive.~~

~~Anomalous values will be not included in summaries/analysis and will be reviewed and flagged by the Clinical Pharmacologist.~~

There are different levels of classification for ADAb status, the definitions are as follows:

~~For each participant an overall ADAb status will be derived:~~

- ~~Overall Positive is defined as having at least one value that is confirmed positive during the treatment period.~~
- ~~Overall Negative is defined as having no values that are confirmed positive at any time in the treatment period.~~

~~The treatment period does not include Baseline/pre-treatment samples or SFU.~~

~~Furthermore, the following subcategories for each subject will be derived:~~

- **Pre ADAb negative – treatment-emergent ADAb negative (Category 1): Includes study participants who are anti-bimekizumab antibody negative at Baseline and anti-bimekizumab antibody negative at all sampling points during the period of interest (one post-Baseline missing/inconclusive sample is allowed for subjects with pre- anti-bimekizumab antibody negative sample). This group also includes study participants who have a missing or inconclusive sample (either missing or inconclusive or insufficient volume) at Baseline (ie, pre-treatment) with all post-Baseline samples as ADAb negative.**
- **Pre ADAb negative – treatment-emergent ADAb positive (Category 2): Includes study participants who are ADAb negative at Baseline and ADAb positive at any sampling points post-Baseline during the period of interest. This group also includes study**

participants who have a missing sample (either missing or insufficient volume) at Baseline (ie, pre-treatment) with 1 or more post-Baseline samples as ADAb positive.

- **Pre ADAb positive – treatment-emergent reduced ADAb (Category 3):** Includes study participants who are ADAb positive at Baseline, and ADAb negative at all sampling points post-Baseline during the period of interest.
- **Pre ADAb positive – treatment-emergent unaffected ADAb positive (Category 4):** Includes study participants who are ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with titer values of the same magnitude as Baseline (ie, less than a predefined fold difference from the Baseline titer).
- **For this analysis, this is set at an increase of less than the validated Minimum Significant Ratio (MSR) of 2.07-fold from Baseline.**
- **Pre ADAb positive – treatment-emergent ADAb boosted positive (Category 5):** Includes study participants who ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with increased titer values compared to Baseline (equal to or greater than a predefined fold difference increase from Baseline titer which will be defined within the validation of the assay).
 - **For this analysis, this is set at an increase equal to or greater than the validated MSR of 2.07-fold from Baseline.**
 - **Note:** for any study participant who is ADAb positive at Baseline and ADAb positive at a post-Baseline time point during the period of interest, but for whom titers are not available to determine treatment unaffected or treatment boosted status, the study participant will be considered as treatment boosted, assuming no other samples are available.
- **ADAb Inconclusive (Category 6):** Includes study participants who have an ADAb positive Baseline (pre-treatment) sample and some post-Baseline samples during the period of interest are missing or inconclusive, while other post-Baseline samples are ADAb negative.
- **Total treatment-emergent ADAb positivity (Category 7 [Categories 2 and 5 combined]):** Includes study participants who are pre ADAb negative – treatment-emergent ADAb positive (Category 2) and pre ADAb positive – treatment boosted ADAb positive (Category 5).
- **Total prevalence of pre- ADAb positivity (Category 8 [Categories 3, 4, 5 and 6 combined]):** Study participants that are tested ADAb positive at Baseline.
- **Missing:** Includes study participants who are ADAb negative, missing, or inconclusive at Baseline with some post-Baseline samples as missing or inconclusive, and other samples as ADAb negative.

Derivation for above classification will be different for the interim analysis and the final analysis. SFU data will be considered only for the final data analysis. That is, each instance of “excluding SFU” in the categories above, should be changed to “including SFU.”

In the case that a sample is collected 1 or more days following the scheduled visit date in which the drug was administered, the **ADAb anti-bimekizumab antibody** results for that sample will be associated with the scheduled visit and summarized accordingly. Such samples will also be considered when **ADAb anti-bimekizumab antibody** results are summarized over a given study period.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by participants. All analyses will be run on the **AMS Active Medication Set**, unless specified otherwise.

- Summary of **ADAb anti-bimekizumab antibody** status overall and by each visit separated by treatment group
- Summary of the time-point of the first occurrence of **ADAb anti-bimekizumab antibody** positivity during the treatment period by treatment group. A plot of the titer by time to first **ADAb anti-bimekizumab antibody** positivity will be prepared.
- All individual participant-level **ADAb anti-bimekizumab antibody** results will be listed.
- The number and percentage of participants in each of the **8 ADAb anti-bimekizumab antibody** categories during the treatment period by treatment group, with an additional category combining participants in categories 2 and 5, summarized as total treatment emergent. In addition, the count and percentage of participants who are pre anti-bimekizumab positive will be calculated (this is the sum of categories 3, 4, and 5).
- The prevalence of immunogenicity, separated by treatment group, and defined subcategory, will be reported by visit, defined as (cumulative) proportion of participants having confirmed positive **ADAb anti-bimekizumab antibody** samples at any visit up to and including that visit. Missing samples will not be included in the denominator.
- The time to achieving treatment-emergent **ADAb anti-bimekizumab antibody** positivity, separated by treatment group and **defined subcategory** sub-categories 2 & 5 above, will be analyzed based on Kaplan-Meier methods. Participants will be considered to have an event at the time point at which treatment emergent **ADAb anti-bimekizumab antibody** positive is first achieved (taking the MSR into consideration for sub-category 5). Participants classified as treatment-emergent **ADAb anti-bimekizumab antibody** negative will be censored at the time of the last available **ADAb anti-bimekizumab antibody** result.
- A summary of HiSCR₅₀ responders at Week 16, separated by treatment group, as a function of ADAb titer will be presented graphically. This will be repeated for HiSCR₇₅ responders.
- Individual plots of plasma bimekizumab concentrations/ **ADAb anti-bimekizumab antibody** titer both plotted on the Y-axes by visit (x-axis) for the full treatment period (excluding SFU for interim analyses and including SFU for final analyses) will be presented for participants with and without HiSCR₅₀ response at Week 16.
- Spaghetti plots of ADAb titer (y-axis) by visit (x-axis), separated by treatment group for all **ADAb anti-bimekizumab antibody** positive participants, including Baseline positive participants.

- Box plots of ADAb titer (logscale) by time to first ADAb positivity by treatment group.

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, 2 categories will be used:

- ~~ADAb Anti-bimekizumab antibody~~ positive – This is defined as participants who have ~~ADAb anti-bimekizumab antibody~~ levels above the specified cut point on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- ~~ADAb Anti-bimekizumab antibody~~ negative – Participants who are not defined as anti-bimekizumab positive (as described above) will be defined as ~~ADAb anti-bimekizumab antibody~~ negative.

The groups for defining ~~ADAb anti-bimekizumab antibody~~ status for safety subgroup analyses are as follows:

- AEs starting before first ~~ADAb anti-bimekizumab antibody~~ positive result
- AEs starting on or after first ~~ADAb anti-bimekizumab antibody~~ positive result
- AEs for participants who were always ~~ADAb anti-bimekizumab antibody~~ negative

Change #57

Section 10.1.1 Exposure during the Initial Treatment Period

This section was split into 2 subsection 10.1.1.1 and 10.1.1.2 for exposure duration (days) and time at risk (days), respectively.

Change #58

Section 10.1.1.1 Study medication duration (days)

The following text was updated:

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Period + 14 days.

Note: The use of 14 days assumes a Q2W dosing interval (bimekizumab 320mg Q2W and placebo). For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Initial Treatment Period – Date of first dose in the Initial Period + 28 days).

Note: If date of last dose in the Initial Treatment Period + 14 days (or ~~date of last bimekizumab dose in the Initial Treatment Period~~ + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Treatment Period + 1.

Change #59

Section 10.1.1.2 Time at risk (days)

The section was updated:

Definitions for time at risk (days) are provided as follows:

~~For participants who permanently discontinue study treatment:~~

- ~~• Date of last dose – date of first dose + 14 days~~

~~The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Initial Treatment Period – date of first dose in the Initial Period + 28 days).~~

~~Note: If date of last dose + 14 days (or date of last dose of bimekizumab + 28 days for Q4W dosing) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:~~

~~— Final visit date (including PEOT, but not including SFU) – date of first dose + 1.~~

Time at risk (days)

- For participants who complete the Week 16 visit and continue to the Maintenance Treatment Period:
 - Date of first dose in the Maintenance Treatment Period – Date of first dose in the Initial Period + 1.
- For participants who discontinue on or prior to the final visit of the Initial Period, use the minimum of the following:
 - **Date of last dose in the Initial Treatment Period – Date of first dose in the Initial Treatment Period + 141**
 - The total number of days in the Initial Treatment Period (112 days). For AEs that emerged after 112 days but still within the 140 days window, those AEs would be classified as TEAE, but will be excluded from the output based on the Initial Treatment Period. However, these AEs will be included in the AE summaries for Maintenance Treatment Period.
 - Date of last clinical contact – Date of first dose in the Initial Treatment Period + 1.
- For participants who die prior to the final visit of the Initial Treatment Period: Date of death – date of first dose in the Initial Period + 1.

Change #60

Section 10.1.2 Exposure during the Maintenance Treatment Period

This section was split into 2 subsection 10.1.2.1 and 10.1.2.2 for exposure duration (days) and time at risk (days), respectively.

Change #61

Section 10.1.2.1 Study medication duration (days)

The section was updated:

Definitions for study medication duration (days) are provided as follows:

- Date of last dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 14 days.

The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W in the Maintenance Treatment Period, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 28 days).

Note: If date of last dose in the Maintenance Treatment Period + 14 days (or ~~date of last bimekizumab dose in the Maintenance Treatment Period~~ + 28 days in the case of Q4W dosing) extends to a date beyond the final visit date of the Maintenance Treatment Period (not including SFU), then this calculation reverts to:

- Final visit date of the Maintenance Treatment Period (not including SFU) – date of first dose in the Maintenance Treatment Period + 1.
- ~~Note:~~ For participants who die during the Maintenance Treatment Period, then this calculation reverts to:
 - Date of death – Date of first dose in the Maintenance Treatment Period + 1.

~~For participants who permanently discontinue study treatment:~~

- ~~Date of last dose – date of first dose + 14 days~~

~~The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W in the Maintenance Treatment Period, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose in the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 28 days).~~

~~Note: If date of last dose + 14 days (or date of last dose of bimekizumab + 28 days for Q4W dosing) extends to a date beyond the final visit date (including PEOT, but not including SFU), then this calculation reverts to:~~

- ~~– Final visit date (including PEOT, but not including SFU) – date of first dose + 1.~~

Change #62

Section 10.1.2.2 Time at risk (days)

The text was updated:

Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study): ~~Final visit date of the Maintenance Treatment Period – date of first dose in the Maintenance Treatment Period + 1.~~
 - **Date of last visit of the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + 1.**
- For participants who die prior to the final visit of the Maintenance Treatment Period: ~~Date of death – date of first dose in the Maintenance Period + 1.~~
 - **Date of death – Date of first dose in the Maintenance Period + 1.**

- For all other participants, use the minimum of the following:
 - Date of last dose in the Maintenance Treatment Period – Date of first dose in the Maintenance Treatment Period + ~~141~~40 days.

Change #63

Section 10.1.3 Exposure during the Initial and Maintenance Treatment Period

This section was split into 2 subsection 10.1.3.1 and 10.1.3.2 for exposure duration (days) and time at risk (days), respectively.

Change #64

Section 10.1.3.1 Study medication duration (days)

The section was updated:

Definitions for study medication duration (days) are provided as follows:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Period.

Note: The algorithms for calculating these durations are specified in Section 10.1.1.1 and Section 10.1.2.1.

Note: If date of last dose in the Initial Treatment Period + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:

- Sum of study medication durations from the individual Initial and Maintenance Treatment Periods - 1.

~~For participants who do not switch study treatments:~~

- ~~• Date of last dose – Date of first dose + 14 days.~~

~~The use of 14 days assumes a Q2W dosing interval. For participants randomized to bimekizumab 320mg Q4W, this will be adjusted based on the dosing interval (eg, use Date of last bimekizumab dose – date of first dose + 28 days).~~

~~Note: If date of last dose + 14 days (or date of last bimekizumab dose in the Maintenance Treatment Period + 28 days in the case of Q4W dosing) extends to a date beyond the final visit date (including PEOT, not including SFU), then this calculation reverts to:~~

- ~~– Final visit date (including PEOT, not including SFU) – Date of first dose + 1.~~

- ~~• For participants who die, if date of last dose + 14 days (or + 28 days in the case of Q4W dosing) extends to a date beyond the date of death, then this calculation reverts to:~~

- ~~– Date of death – Date of first dose + 1.~~

~~For participants who switch study treatments (between Initial and Maintenance Treatment Periods):~~

- ~~• Initial Treatment Period (attributed to initially randomized treatment):~~

- ~~– Date of last dose in the Initial Period – Date of first dose in the Initial Period + 14 days.~~

~~Note: Participants who switch study treatments are on a Q2W dosing schedule for the Initial Treatment Period.~~

~~Note: If date of last dose in the Initial Treatment Period + 14 days extends to a date beyond the date of first dose in the Maintenance Treatment Period, then this calculation reverts to:~~

~~— Date of first dose in the Maintenance Treatment Period — Date of first dose in the Initial Period + 1.~~

- ~~• Maintenance Treatment Period (attributed to the treatment initiated in the Maintenance Treatment Period):~~

~~— Use the study medication duration algorithm specified for the Maintenance Treatment Period in Section 10.1.2.1.~~

Change #65

Section 10.1.3.2 Time at risk (days)

This section was updated:

~~For participants who do not switch study treatments:~~ Definitions for time at risk (days) are provided as follows:

- For participants who complete the Maintenance Treatment Period as planned and continue into an extension study (and, therefore, do not have the SFU visit in the feeder study):
 - Final visit date – Date of first dose + 1.
- For participants who die prior to the final visit:
 - Date of death – Date of first dose in the + 1.
- For all other participants, use the minimum of the following:
 - Date of last dose – Date of first dose + 141 days.
 - Date of last clinical contact – Date of first dose + 1.

Note: This group could include participants who discontinue early, participants who complete the Maintenance Treatment Period as scheduled but choose not to continue into an extension study, or participants who are ongoing in the SFU period at the time of the data snapshot (in the case of the interim analysis).

~~For participants who switch study treatments (between Initial and Maintenance Treatment Periods):~~

- ~~• Initial Treatment Period (attributed to initially randomized treatment):~~

~~— Date of first dose in the Maintenance Treatment Period — Date of first dose in the Initial Period + 1.~~

~~(Note: This assumes that anyone in this category has completed the Initial Treatment Period and doses [with a new study treatment] in the Maintenance Treatment Period.)~~

- ~~• Maintenance Treatment Period (attributed to the treatment initiated in the Maintenance Treatment Period):~~

~~Use the time at risk algorithm specified for the Maintenance Treatment Period in Section 10.1.2.2.~~

Change #66

Section 10.2.1 Data considerations

The following sentence was added:

If the intensity of an AE is unknown, it is considered as severe. If the relationship to study drug is missing, it is considered as related. **Note that if the seriousness of an adverse event is unknown, every attempt should be made to resolve this prior to a snapshot for an interim analysis or database lock; in the exceptional case that the seriousness of an adverse event is still missing then no imputation should be applied for this characteristic.**

Change #67

Section 10.2.2 AE summaries

The following text was deleted:

The following summaries will be provided by treatment group for the Initial Treatment Period, ~~Maintenance Treatment Period~~, and the Initial and Maintenance Treatment Period combined based on the SS, ~~MS~~, and AMS respectively.

The following summaries were added:

- Incidence of TEAEs – Suspected and Confirmed COVID-19 cases by SOC, HLT and PT

Suspected and confirmed COVID-19 cases will be identified with the preferred terms “Corona virus infection” or “Corona virus test positive”.

The following subset of tables will also be presented for the Maintenance Treatment Period using the MS:

- Incidence of TEAEs – Overview
- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of Serious TEAEs per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Discontinuation per 100 subject years by SOC, HLT, and PT
- Incidence of TEAEs Leading to Death by SOC, HLT, and PT

Change #68

Section 10.2.3.1 Infections (serious, opportunistic, fungal and TB)

The following text was updated:

- **Incidence of Fungal Infection TEAEs per 100 subject years by SOC, HLT and PT**

Fungal infections will be summarized in a stand-alone table. The table will include all TEAEs (serious and non-serious) that code into the **High Level Group Term (HLGT)** ~~HLT~~ “Fungal infectious disorders”

- **Incidence of Opportunistic Infection TEAEs per 100 subject years by SOC, HLT and PT**

Opportunistic infections (including tuberculosis) will be summarized in a stand-alone table. The table will include all TEAEs identified using UCB-defined search criteria. (refer to Excel spreadsheet on “OI – MedDRA v19.0.xlsx” in “Bimekizumab Safety Topics of Interest.docx”).

Change #69

Section 10.2.3.3 Major adverse cardiac event

The following sentence was added:

A separate table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type (24 total), the individual PTs that fall within each event type will be summarized. **The other 10 MACE events not listed in the table are described in the adjudication committee charter.**

Change #70

Section 10.2.3.5 Suicidal ideation and behaviors

The following paragraph was added:

A separate table will present the adjudicated neuropsychiatric events by type. For each neuropsychiatric event type (6 total), the individual PTs which fall within each event type will be summarized. It will include events adjudicated as SIB and events adjudicated as non-suicidal. Note that the event type Suicidal ideation can be classified as either SIB or non-suicidal.

Change #71

Section 10.2.3.6 Inflammatory bowel disease

This section was updated:

An external inflammatory bowel disease (IBD) adjudication committee will evaluate potential IBD events and will classify each one as follows:

- Event Type Code 1: Possible IBD – Crohn’s Disease
- Event Type Code 2: Probable IBD – Crohn’s Disease
- Event Type Code 3: Definite IBD – Crohn’s Disease
- Event Type Code 4: Possible IBD – Ulcerative Colitis
- Event Type Code 5: Probable IBD – Ulcerative Colitis
- Event Type Code 6: Definite IBD – Ulcerative Colitis
- Event Type Code 7: Possible IBD – Unclassified
- Event Type Code 8: Probable IBD – Unclassified
- Event Type Code 9: Definite IBD – Unclassified
- Event Type Code 10: Symptoms not consistent with IBD
- **Event Type Code 11: Possible Inflammatory Bowel Disease – Microscopic Colitis**
- **Event Type Code 12: Probable Inflammatory Bowel Disease – Microscopic Colitis**

- **Event Type Code 13: Definite Inflammatory Bowel Disease – Microscopic Colitis**
- **Event Type Code 14: Possible Inflammatory Bowel Disease – no further differentiation possible**
- **Event Type Code 15: Probable Inflammatory Bowel Disease – no further differentiation possible**
- **Event Type Code 16: Definite Inflammatory Bowel Disease – no further differentiation possible**
- **Event Type Code 99: Not enough information to adjudicate**

A table for adjudicated ~~definite~~ IBD events (event type codes **1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15** and **169**) as determined by the adjudication committee will be produced. **It will summarize events determined by the adjudication committee as definite IBD (event type codes 3, 6, 9, 13, and 16), probable IBD (event type codes 2, 5, 8, 12, and 15) and possible IBD (event type codes 1, 4, 7, 11, and 14). Definite and probable IBD will also be aggregated and summarized.** This table will be ~~produced overall~~, as well as stratified by subjects with or without a previous medical history of IBD. Previous medical history of IBD will be determined using the information recorded on the **History of IBD Extra-Articular Assessment at Screening** CRF page ("Does subject have a history of IBD?").

A ~~separate table will present the for~~ adjudicated gastrointestinal ~~probable IBD events by type.~~ **For each gastrointestinal event type (17 total), the individual PTs which fall within each event type will be summarized. It will include events codes 2, 5, and 8) as determined by the adjudication committee as definite IBD probable IBD and will be produced.** This table will be ~~produced overall~~, as well as stratified by participants with or without a previous medical history of IBD.

A table for adjudicated ~~possible IBD. It~~ events (event type codes ~~1, 4, and 7~~) as determined by the adjudication committee will be produced. This table **will also include events determined as Symptoms not consistent with** be produced overall, as well as stratified by participants with or without a previous medical history of **IBD (event type code 10) and Not enough information to adjudicated (event type code 99).**

A listing of all events identified for potential review by the IBD adjudication committee will be produced. This listing will indicate whether each event was escalated to the committee for formal review/adjudication.

A ~~separate table and listing~~ will present the adjudicated IBD events by type. For each IBD event type (event type codes 1 through ~~1640~~ and 99; ~~1744~~-total), the individual PTs which fall within each event type will be ~~listed~~**summarized.**

A third listing will present the individual diagnostic criteria met for each adjudicated IBD event.

Change #72

Section 10.2.3.7 Hypersensitivity (including anaphylaxis)

The following text was updated:

A separate table will be prepared to summarize hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All TEAEs which code to a PT included in the Scope=Narrow search will be included in this table. **In addition, a separate table will be prepared to summarize serious hypersensitivity events, identified using the SMQ “Hypersensitivity (SMQ)”. All serious TEAEs which code to a PT included in the Scope=Narrow search will be included in this table.** An AE glossary table will also be produced to summarize the MedDRA coding for these events. The glossary table will include the following fields: reported term, PT, LLT, HLT, and SOC.

Furthermore, **a separate table will be prepared to summarize injection site reactions, identified using the HLTs: “Administration site reactions NEC” and “Injection site reactions”.**

Change #73

Section 10.3 Clinical laboratory evaluations

The following text was added:

For tables where data are summarized by visit, unscheduled and repeat visits will not be summarized, but these data will be included in listings. For tables where multiple measurements over a period of time are considered (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (**values observed more than 140 days after the last administration of study medication are not considered**). All summaries will be presented in SI units and will be based on observed case values.

CTCAE grading was updated:

Markedly abnormal values are defined as those with a severity of Grade 3 and above based on the CTCAE criteria **Version 4.03**, (U.S. Department of Health and Human Services 2017).

And Table 10-2 was updated:

Table 10–2: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Creatinine ¹ Creatinine	mg/dL	> 3.0 x Baseline or >3.0 x ULN	umol/mole μmol/L	> 3.0 x Baseline or >3.0 x ULN	AH
Glucose	mg/dL	<40 >250	mmol/L	<1.7	AL
				>13.9	AH
Calcium	mg/dL	>12.5 <7.0	mmol/L	>3.1	AH
				<1.75	AL
Magnesium	mg/dL	>3.0 <0.9	mmol/L	>1.23	AH
				<0.4	AL
Potassium	mmol/L	>6.0 <3.0	mmol/L	>6.0	AH
				<3.0	AL

Table 10–2: Definitions of Markedly Abnormal Biochemistry Values

Parameter name	Conventional		Standard		Abnormal Designation
	Unit	Criteria	Unit	Criteria	
Sodium	mmol/L	>155 <130	mmol/L	>155 <130	AH AL
Cholesterol	mg/dL	>400	mmol/L	>10.34	AH

1 The markedly abnormal definitions for creatinine are based on the logical or, if either criterion is met the creatinine value will be designated as abnormal high.

And the following text was added:

- **Total Bilirubin:** >1.5xULN, >2xULN
- **ALP:** >1.5xULN

For any participant with at least one markedly abnormal LFT (AST >3xULN, ALT >3xULN, bilirubin >3xULN, or ALP >1.5xULN) the New Ratio (nR) will be calculated as the ratio of either ALT or AST (whichever is higher) to ALP, all expressed as multiples of their ULN as follows:

- $nR = [\text{maximum}(\text{AST}/\text{ULN or ALT}/\text{ULN})]/(\text{ALP}/\text{ULN})$

Any pDILI will be summarized (all criteria must be met at the same assessment):

- (AST or ALT > 3xULN) and Total Bilirubin > 1.5xULN
- (AST or ALT > 3xULN) and Total Bilirubin > 2xULN

In addition, a table will be produced to summarize potential Hy's Law cases. The following definition will be used in that table:

- $[\text{AST} \geq 3\text{xULN or ALT} \geq 3\text{xULN}] \text{ and Total Bilirubin} \geq 2\text{xULN}$ in the absence of ALP $\geq 2\text{xULN}$

In order to meet the above **potential Hy's Law** criteria, a participant must experience the elevation in bilirubin and ALT or AST (and the absence of ALP elevation, if applicable) at the same assessment. For example, a participant who experiences a ≥ 2 x ULN elevation of bilirubin at one visit and a $\geq 3\text{xULN}$ elevation in ALT (or AST) at a subsequent visit has not fulfilled the Hy's Law criteria.

Potential hepatotoxicity (meeting one of the PDILI or Hy's Law laboratory criteria at least once) will be considered with and without symptoms potentially associated with hepatitis or hypersensitivity according to the investigator (reported on the Symptoms of Hepatitis and Hypersensitivity CRF page).

Change #74

Section 10.4.1 Vital signs

The following text was added:

Unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

Change #75

Section 10.4.3 Other safety endpoints

The following text was added:

For by-visit summaries, unscheduled and repeat visits will not be summarized, but these data will be included in listings. By-visit tables should include the SFU visit. Summaries over a period of time (as in shift tables), unscheduled and repeat visits will be considered as long as they were collected in the period being summarized (values observed more than 140 days after the last administration of study medication are not considered).

Change #76

Section 10.4.3.2 Columbia-Suicide Severity Rate Scale (C-SSRS)

The following text was updated:

The incidence of participants with suicidal ideation, suicidal behavior, suicidal ideation or behavior, and self-injurious behavior will be summarized **for the Initial Treatment Period and the combined Initial and Maintenance Treatment Period** by treatment group ~~for each treatment.~~

Change #77

Section 10.3.4.5 Patient Health Questionnaire (PHQ)-9 scores

The following text was updated:

~~A In addition, a categorical summary of the absolute and change from Baseline value scores will be presented by treatment group and visit.~~

~~The percentage of study participants with scores below a corresponding shift table. The following categories will be presented: 0-4; 5, between 5 and 9, between 10 and 14, between 15 and 19, and greater than or equal to 20 in PHQ-9 will be summarized as a shift from Baseline by visit and treatment group based on observed values; ≥ 20 .~~

The percentage of study participants with scores ≥ 15 at any post-Baseline visit and the number and percentage of study participants with scores ≥ 20 at any post-Baseline visit will be summarized by treatment group based on observed values. This summary will also include the percentage of study participants with increase from baseline ≥ 5 at any post-Baseline visit.

~~Different to other safety variables, PHQ-9 will be summarized using the MCMC/monotone regression approach described for continuous variables.~~

The number and percentage of participants that complete the PHQ-9 will be calculated for each visit by treatment group. The percentage will be based on the number of participants in the RS. ~~(or MS, as appropriate).~~ The summary will be repeated where the percentage is based on the subset of participants who complete each particular visit.

Change #78

Section 11 References

The following reference was updated:

Common Terminology Criteria for Adverse Events (CTCAE); Version **4.0** June **2010**
2017. U.S. Department of Health and Human Services

Change #79

Section 12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The algorithm for identifying anaphylaxis was updated:

The SMQ *Anaphylactic reaction* consists of three parts:

- A **narrow search** containing PTs that represent core anaphylactic reaction terms (Category A – core anaphylactic reaction terms)

Anaphylactic reaction
Anaphylactic shock
Anaphylactic transfusion reaction
Anaphylactoid reaction
Anaphylactoid shock
Circulatory collapse
Dialysis membrane reaction
Kounis syndrome
Shock
Shock symptom
Type I hypersensitivity

- A **broad search** that contains additional terms that are added to those included in the narrow search. These additional terms are signs and symptoms possibly indicative of anaphylactic reaction and categorized in B, C or D

Category B (Upper Airway/Respiratory Terms)

Acute respiratory failure	Nasal obstruction
Asthma	Oedema mouth
Bronchial oedema	Oropharyngeal spasm
Bronchospasm	Oropharyngeal swelling
Cardio-respiratory distress	Respiratory arrest
Chest discomfort	Respiratory distress
Choking	Respiratory failure
Choking sensation	Reversible airways obstruction

Circumoral oedema	Sensation of foreign body
Cough	Sneezing
Cyanosis	Stridor
Dyspnoea	Swollen tongue
Hyperventilation	Tachypnoea
Irregular breathing	Throat tightness
Laryngeal dyspnoea	Tongue oedema
Laryngeal oedema	Tracheal obstruction
Laryngospasm	Tracheal oedema
Laryngotracheal oedema	Upper airway obstruction
Mouth swelling	Wheezing

▪ **Category C (Angioedema/Urticaria/Pruritus/Flush terms)**

Allergic oedema	Oedema
Angioedema	Periorbital oedema
Erythema	Pruritus
Eye oedema	Pruritus allergic
Eye pruritus	Pruritus generalised
Eye swelling	Rash
Eyelid oedema	Rash erythematous
Face oedema	Rash generalised
Flushing	Rash pruritic
Generalised erythema	Skin swelling
Injection site urticaria	Swelling
Lip oedema	Swelling face
Lip swelling	Urticaria
Nodular rash	Urticaria papular
Ocular hyperaemia	

▪ **Category D (Cardiovascular/Hypotension terms)**

Blood pressure decreased
Blood pressure diastolic decreased
Blood pressure systolic decreased
Cardiac arrest
Cardio-respiratory arrest
Cardiovascular insufficiency

Diastolic hypotension
Hypotension

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after, and for scenarios where two events must have been reported, both events must have occurred within one day of each other:
 - A narrow term or a term from Category A;
 - A term from Category B - (Upper Airway/Respiratory) AND a term from Category C - (Angioedema/Urticaria/Pruritus/Flush);
 - A term from Category D - (Cardiovascular/Hypotension) AND [a term from Category B - (Upper Airway/Respiratory) OR a term from Category C - (Angioedema/Urticaria/Pruritus/Flush)]

Change #80**Section 912.2 Appendix B: Definition of CTCAE grades**

Table 12-1 was updated:

Table 12–1: Definitions of CTCAE grades by biochemistry parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Creatinine ¹	High	umol/L mmol/L	>1-1.5x Baseline or >ULN-1.5 x ULN	>1.5-3.0x Baseline or >1.5 – 3.0 x ULN	>3.0x Baseline or >3.0 – 6.0 x ULN	>6.0 x ULN
Sodium	Low	mmol/L	130-<LLN	N/A	120-<130	<120
Sodium	High	mmol/L	>ULN-150	>150-155	>155-160	>160
Potassium ²	Low	mmol/L	3.0-<LLN	3.0-<LLN	2.5-<3.0	<2.5
Potassium	High	mmol/L	>ULN-5.5	>5.5-6.0	>6.0-7.0	>7.0
Calcium	Low	mmol/L	2.0-<LLN	1.75-<2.0	1.5-<1.75	<1.5
Calcium	High	mmol/L	>ULN-2.9	>2.9-3.1	>3.1-3.4	>3.4
Magnesium	Low	mmol/L	0.5-<LLN	0.4-<0.5	0.3-<0.4	<0.3
Magnesium	High	mmol/L	>ULN-1.23	N/A	>1.23-3.30	>3.30
Cholesterol	High	mmol/L	>ULN-7.75	>7.75-10.34	>10.34- 12.82	>12.82

1 The CTCAE Grade definitions for creatinine are based on the logical or and the highest applicable CTCAE grade should be assigned to a creatine value.

2 The decreased potassium criterion of 3.0-<LLN is specified for both CTCAE Grade 1 and Grade 2; values meeting this criterion will be counted as Grade 2.

And Table 12-2 was updated:

Table 12–2: Definitions of CTCAE grades by hematology parameter

Parameter	Definition	Unit	Grade 1	Grade 2	Grade 3	Grade 4
Hemoglobin	Low	g/L	100-<LLN	80-<100	<80	N/A
Hemoglobin ¹	High	g/L	>0-20 above ULN or >0-20 above Baseline if Baseline is above ULN	>20-40 above ULN or >20-40 above Baseline if Baseline is above ULN	>40 above ULN or >40 above Baseline if Baseline is above ULN	N/A
Platelets	Low	10 ⁹ /L	75-<LLN	50-<75	25-<50	<25
WBC	Low	10 ⁹ /L	3-<LLN	2-<3	1-<2	<1
WBC	High	10 ⁹ /L	N/A	N/A	>100	N/A
Lymphocytes	Low	10 ⁹ /L	0.8-<LLN	0.5-<0.8	0.2-<0.5	<0.2
Lymphocytes	High	10 ⁹ /L	N/A	>4-20	>20	N/A
Neutrophils	Low	10 ⁹ /L	1.5-<LLN	1.0-<1.5	0.5-<1.0	<0.5

LLN=lower limit of normal; N/A=not applicable; ULN=upper limit of normal, WBC=white blood cells

1 The CTCAE Grade definitions to be applied are dependent on the Baseline hemoglobin value. If the baseline value is > ULN then the criteria relative to Baseline is applicable, otherwise the criteria relative to ULN is applicable.

13.2 Amendment 2

Rationale for the amendment

The main purposes of this amendment were:

- General update to analyses to align with protocol amendment 4.
- Procedural clarifications from discussions and feedback provided at meetings
- Update to align with the bimekizumab program standards and safety topics of interest

Modifications and changes

Global Changes

The following changes were made throughout the SAP:

- Typos and formatting were updated throughout the document
- HSSDD worst pain and average pain were updated to worst skin pain and average skin pain, respectively

Specific changes

In addition to the global changes, the following specific changes have been made (typos such as missing spaces or redundant spaces are not listed):

Change #1

List of Abbreviations

The following abbreviations have been added:

CFB	change from Baseline
CV-CAC	Cardiovascular Event Adjudication Committee
eCDF	empirical cumulative distribution function
IBD-CAC	Inflammatory Bowel Disease Adjudication Committee

Change #2

Section 1 Introduction

The protocols were updated:

The SAP is based on the Protocol Amendment **4 3, 6 May 2022** ~~9 February 2021~~ and the Japan-specific amendment **4.1 3-1, 11 February 2021**.

Change #3

Section 2.2 Study endpoints

The following text was deleted:

The endpoints based on HS Symptom Daily Diary (HSSDD) and Hidradenitis Suppurativa Symptom Questionnaire (HSSQ) pain responses are based on the current definitions, which are continuous. It is anticipated that a responder (binary) endpoint will be defined for the HSSDD and HSSQ pain items as well as other symptom items prior to database lock and unblinding, based on separate ongoing, blinded, psychometric analyses aiming to determine threshold for within-patient clinically meaningful improvement.

The below HSSDD and HSSQ pain response endpoints and analyses will be adjusted accordingly in a future SAP amendment.

Change #4

Section 2.2.1.2 Secondary efficacy endpoints

The following bullet was added:

- Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline

Change #5

Section 2.2.1.3 Other efficacy endpoints

The following bullets were updated:

- ~~Skin pPain response status~~ain response, as assessed by the “worst pain” item in the HSSDD, defined as an improvement from baseline in the weekly worst skin pain score of at least 3 units
- **Skin pain response based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) among study participants with a score of ≥ 3 at Baseline**

- **Skin pain response** ~~Response in HS Skin Pain (11-point numeric rating scale) assessed by the HSSDD at Week 16~~ **Response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly worst skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline**
- **Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HSSDD weekly average skin pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline**
- ~~HS Symptom Questionnaire (HSSQ) response (at least a 3-unit reduction from Baseline in worst HS Skin Pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline assessed HSSQ by the HSSDD in the Initial Treatment Period, and assessed by the HSSQ in the Maintenance Treatment Period~~ **Skin pain response (at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) assessed by the HSSQ among study participants with a score of ≥ 3 at Baseline**
- Absolute change from Baseline in DLQI Total Score
- DLQI Total Score of 0 or 1
- Minimum clinically important difference (MCID) (improvement from Baseline of 4 or more) in the DLQI Total Score among study participants with a Baseline score of at least 4
- Absolute change from Baseline in Hidradenitis Suppurativa Quality of Life (HiSQOL) domain scores (symptoms, psychosocial, activities and adaptations) and Total score
- Patient Global Impression of HS Severity (PGI-S-HS)
- Patient Global Impression of Change of HS Severity (PGI-C-HS)
- Patient Global Impression of Severity of Skin Pain (PGI-S-SP)
- Patient Global Impression of Change of Skin Pain (PGI-C-SP)
- Absolute change from Baseline in each of the other HS Symptoms - itch, drainage or oozing of HS lesions, and smell or odor.
- **Responders Response** on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor

Change #6

Section 2.4 Determination of sample size

The power to detect a statistically significant difference for each of the endpoints are shown in [Table 2-1](#). Notably, with a 2-sided significance level of 0.025, the sample size of 140:70 provides 73% power for detecting at least a difference of 1.5 (bimekizumab Q4W vs placebo) for the Worst **Skin Pain change from Baseline (CFB)** endpoint.

Given the high level of power for each of the primary and key secondary endpoints at the 0.025 significance level for the Q2W comparison (power ≥ 0.89), and per the alpha spending strategy, there is a high likelihood that the Q4W comparison of Worst **Skin Pain CFB** vs placebo ~~for Worst Pain change from Baseline~~ will be allowed to be tested against the 0.05 level of

significance. The power for this latter test is 81%. The sample size is thus ultimately driven by the Worst Skin Pain CFB endpoint. Furthermore, the randomization ratio of 2:2:2:1 has been chosen to provide study participants with a high probability (6/7 ~ 86%) of being randomized to active study drug.

After randomization for this study was complete, an additional endpoint to assess Worst Skin Pain response was included in the sequential testing procedure. This additional endpoint is based on the threshold for clinically meaningful change and is defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline. Note that the power calculations reported in Table 2-1 for this endpoint are based on the sample size that was initially driven by the Worst Pain CFB endpoint as described above. With a 2-sided significance level of 0.025, the sample size of 104:52 in the subset of participants reporting Baseline HSSDD worst skin pain score at or beyond the threshold for clinically meaningful change (ie, Baseline HSSDD ≥ 3) provides 53% power for detecting a statistically significant difference between bimekizumab Q4W and placebo in the proportion of Worst Pain responders.

Note that the power at the 0.025 level of significance associated with this endpoint for the comparison between bimekizumab Q2W and placebo is 95%. The Q4W comparison of Worst Skin Pain response vs placebo against the 0.05 level of significance is therefore likely, and the power at this significance level is 65%. Given the strength of the power for the bimekizumab Q2W arm vs placebo across endpoints, it is considered acceptable to have a relatively low power for this final endpoint in the testing sequence for the bimekizumab Q4W treatment arm vs placebo.

Change #7

Section 2.4 Determination of sample size

Table 2-1 was updated:

Table 2-1: Power calculation assumptions and methods

Endpoint	Power $\alpha = 0.025$, 2-sided		Assumptions		
	Q2W	Q4W	Week 16 Bimekizumab Q2W N=280 ^a	Week 16 Bimekizumab Q4W N=140	Week 16 Placebo N=70
HiSCR ₅₀	0.99	0.90	Proportion responders=0.60	Proportion responders=0.50	Proportion responders=0.25
HiSCR ₇₅	0.99	0.98	Proportion responders=0.45	Proportion responders=0.35	Proportion responders=0.10
Flare	0.99	0.99	Proportion of participants with flare by Week 16=0.09	Proportion of participants with flare by Week 16=0.19	Proportion of participants with flare by Week 16=0.52
DLQI	0.99	0.96	Mean CFB=-5.4; SD=6.8	Mean CFB=-4.8; SD=6.8	Mean CFB=-0.8; SD=6.6

Table 2–1: Power calculation assumptions and methods

Worst Skin Pain CFB ^b	0.89	0.73	Mean CFB=-2.2; SD=3.2	Mean CFB=-2.0; SD=3.2	Mean CFB=-0.5; SD=3.7
Worst Pain Response ^c	0.95	0.53	Proportion responders=0.53	Proportion responders=0.43	Proportion responders=0.23

CFB=change from Baseline; Q2W=every 2 weeks; Q4W=every 4 weeks; SD=standard deviation

Note: Estimates for HS0004 are based on Week 12 data from the HS0001 study.

^a Pooled Q2W at Week 16 from Q2W/Q2W and Q2W/Q4W arms

^b Within-participant average of Worst Skin Pain according to 24-hour recall.

^c Assumes N=208, 104, 52 in Q2W, Q4W, and placebo, respectively, to account for Worst Skin Pain score at or above 3 (ie, the threshold for clinically meaningful change from Baseline).

Change #8

Section 3.1 General presentation of summaries and analyses

The following text was added:

Per protocol, visit windows are ± 3 days from the date of first dose. The 20-week SFU Visit window is ± 7 days from the date of the final dose. All by-visit summaries will contain nominal (ie, scheduled) visits only. Unscheduled visits will not be mapped to scheduled visits except for assessments that occur within a 3-day time window of a scheduled visit. In that case, the assessment will be mapped to the corresponding scheduled visit and will be used for the analysis. This will only occur for selected vendor data. **The only exception to this rule is for unscheduled assessments that occur up to 3 days after the Baseline visit. These unscheduled visits will remain as unscheduled as the Baseline assessment cannot be after the first dose of study drug administration. See Section 3.3 for more details on the definition of Baseline values.**

Change #9

Section 3.5.8 Pharmacokinetics Per-Protocol Set

The following text was deleted:

The Pharmacokinetics Per-Protocol Set (PK-PPS) will consist of study participants who received at least 1 full dose of bimekizumab and provided at least 1 quantifiable plasma concentration post-dose without important protocol deviations that would affect the PK. ~~The Pharmacokinetics Per-Protocol Set is defined separately for each of the treatment periods (ie, separately for the Initial Treatment Period and the Maintenance Treatment Period).~~

Change #10

Section 3.10 Changes to protocol-defined analyses

The following text was updated:

The following other efficacy endpoints are included in the protocol but will not be included as part of the analysis:

- **Responders-Response** on other HS Symptoms (11-point numeric rating scale) - itch, drainage or oozing of HS lesions, and smell or odor
- Flare (defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Week 16) by Week 48

The calculation of nominal p-values has been added for selected efficacy endpoints. **These nominal p-values are not controlled for multiplicity and should not be used to declare statistical significance.**

The protocol defines the PK-PPS separately by period, but there will only be one PK-PPS for the overall study.

Change #11

Section 4.1 Adjustments for covariates

The following sentence was updated:

The Worst Skin Pain **secondary endpoints (change from Baseline continuous secondary endpoint and pain response binary endpoint)** will also include analgesic use as a covariate.

Change #12

Section 4.2.1.2 Handling of missing data for the secondary efficacy endpoints

The following paragraph was updated:

For secondary continuous efficacy endpoints, MI-MCMC/monotone regression is the primary method for imputing missing data, regardless of whether the missing data are preceded by an intercurrent event. That is, if an intercurrent event occurs on or before a visit, the result for that visit will be treated as missing and **then imputed with the missing data**. If the imputation model cannot converge, last observation carried forward (LOCF) will be used. The OC method will be performed as a sensitivity analysis.

Change #13

Section 42.2.1 MI – MCMC/Monotone Regression

The imputation rule for HSSDD was updated in Table 4-2.

Table 13–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
Lesion count ^a	0	--	Yes
DLQI total score	0	30	Yes
hs-CRP	LLOQ/2	--	No
HSSDD item score	0	10	NoYes
HSSQ item score	0	10	Yes
HiSQOL symptom status score	0	16	Yes

Table 13–2: Imputation allowable ranges by variable

Variable	Minimum Value	Maximum Value	Integer Values Only
HiSQOL psychosocial impact score	0	20	Yes
HiSQOL impact on physical activities score	0	32	Yes
EQ-5D-3L VAS	0	100	Yes
WPAI dimension scores	0	100	No for variables: “Percent work time missed due to problem” and “Percent overall work impairment due to problem” Yes for variables: “Percent impairment while working due to problem” and “Percent activity impairment due to problem”. These two variables can only take values that are multiples of 10.

^a Lesion counts will be imputed separately for each lesion type (abscesses, draining tunnels [fistulas/sinus tracts], inflammatory nodules, non-draining tunnels [fistulas/sinus tracts], non-inflammatory nodules, HS scars). The imputed lesion counts will be used to derive the endpoints that are dependent on the lesion count data (eg, HiSCR₅₀).

Change #14

Section 4.2.3 Rationale for estimand

The following text was added to the bullet:

- A composite estimand strategy will be used for the primary analysis of the primary and binary secondary endpoints (HiSCR₅₀, HiSCR₇₅, flare, **HS worst skin pain response**),

Change #15

Section 4.5 Multiple comparisons/multiplicity

Under this framework, each bimekizumab dose of 320mg Q2W and 320mg Q4W will be compared to placebo in the first instance at a familywise error rate of 0.025 ($\alpha/2$).

Simultaneously within each dose, closed testing for the primary and secondary efficacy endpoints will be performed as follows:

1. Step 1: Test HiSCR₅₀ at significance level 0.025.
2. Steps 2 to ~~65~~ – If Step 1 is significant at 0.025 then test each secondary endpoint sequentially in the order shown in [Figure 4-1](#), moving to the next step only if significance achieved at 0.025.
3. In the event that Step ~~65~~ is significant at 0.025 for a given dose, then Steps 1 to 6 will be repeated for the other dose using a significance level of 0.05.

The secondary efficacy variables supporting the primary efficacy variable are listed below, and will be included in the multiplicity adjustment using the analysis methods specified in Section 8.3:

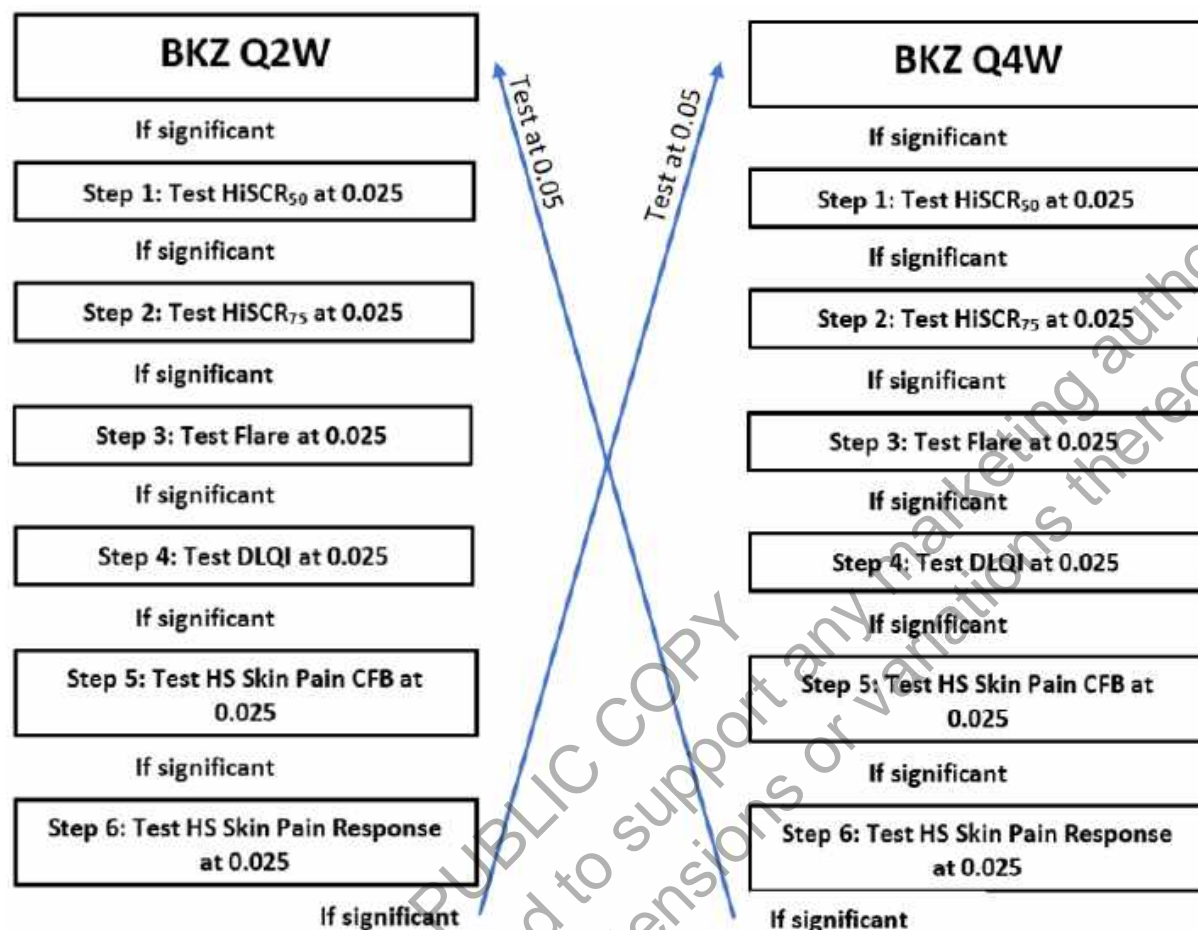
4. Proportion of study participants who achieve HiSCR₇₅ at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
5. Proportion of study participants who experience at least 1 flare by Week 16, with flare defined as a $\geq 25\%$ increase in AN count with an absolute increase in AN count of ≥ 2 relative to Baseline.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
6. CFB in DLQI Total Score at Week 16.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
7. Absolute change from Baseline in Skin Pain Score at Week 16, as assessed by the “worst skin pain” item (11-point numeric rating scale) in the HSSDD.
 - a. bimekizumab 320mg Q2W vs placebo
 - b. bimekizumab 320mg Q4W vs placebo
8. **Skin pain response at Week 16, based on the threshold for clinically meaningful change (defined as at least a 3 point decrease from Baseline in HSSDD weekly worst skin pain score) at Week 16 among study participants with a score of ≥ 3 at Baseline.**
 - a. **bimekizumab 320mg Q2W vs placebo**
 - b. **bimekizumab 320mg Q4W vs placebo**

Change #16

Section 4.5 Multiple comparisons/multiplicity

Figure 4-1 was updated to add the new secondary endpoint:

Figure 13-1: Sequence of testing



AN=abscess and inflammatory nodule; DLQI=Dermatology Life Quality Index; HiSCR₅₀=a 50% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HiSCR₇₅=a 75% reduction in the total AN count with no increase from Baseline in abscess or draining tunnel count; HS=hidradenitis Suppurativa; Q2W=every 2 weeks; Q4W=every 4 weeks

HS skin pain response is tested among study participants with a score of ≥ 3 at Baseline.

Change #17

Section 4.8 Examination of subgroups

The following text was updated:

Subgroup analyses will be performed on the HiSCR₅₀, HiSCR₇₅, ~~and flare~~, **and worst skin pain response** endpoints by visit for the Initial Treatment Period and Maintenance Treatment Period. Additional subgroup analyses will be performed on the CFB in the worst skin pain score as measured by HSSDD and in the DLQI total score through Week 16 as described below.

Along with the tables described, there will be tables for HiSCR₅₀, HiSCR₇₅, ~~and flare~~, **and skin pain response endpoints** which display the response difference and 95% CIs between each bimekizumab dose regimen versus placebo for each of the subgroups at Week 16. Corresponding forest plots will be prepared.

Additionally, the following bullets were clarified:

- Antibody positivity (confirmatory assay: negative or positive. **See Section 9.3.2**)
- **Antihistamines users during the Initial Treatment Period (yes, no) (Section 6.4.2 specifies how participants are classified as antihistamine users) (applicable only to the skin pain response endpoint)**

Change #18

Section 6.4 Prior and concomitant medications

The following text was added:

The number and percentage of study participants with concomitant vaccines for COVID-19 will be summarized by treatment group, overall and by World Health Organization Drug Dictionary Standardized Drug Grouping (SDG), presenting SDG subgroup, and preferred term. The SDG subgroup Vaccines for COVID-19 will be used to identify vaccines for COVID-19 using the narrow scope; this subgroup is divided further into separate subgroups which is the level that will be presented. The number of individual occurrences of the vaccine for COVID-19 will also be summarized.

A listing of concomitant vaccines for COVID-19 will be provided.

Change #19

Section 6.4.2 Classification of participants as analgesic, antihistamine users

The section was updated as follows:

If a participant has taken a new analgesic/increased regimen of analgesic, or taken an antihistamine, on 1 or more days (need not be consecutive) ~~for a given week in a study period (Initial Treatment Period or Maintenance Treatment Period), then for that week period the participant will be classified as an analgesic or antihistamine user, respectively. The week period under consideration is to match the period as defined for the HSSDD week for the Initial Treatment Period or HSSQ week for the Maintenance Treatment Period, based on dates/times of the medications taken. If there is a visit date but no HSSDD available at the visit, then the analgesic/antihistamine user status for that week will be derived based on the visit date. If there is no visit available, then the weekly analgesic/antihistamine user status will default to the analgesic/antihistamine status for the overall study period.~~

New analgesic/increased regimen of analgesic use, regardless of indication, is defined as an analgesic medication with start date on or after the first dose of study medication. Stable analgesics (ie, analgesics which were taken already before randomization) will not be included in this category of analgesic user. This classification will be used ~~to adjust the formal analysis of the Worst Pain secondary endpoints and for selected subgroup analyses.~~

Antihistamine use is identified by considering the ATC classification. This classification is used for analyzing the Worst Itch endpoint and for selected subgroup analyses, by visit, for the Initial Treatment Period and Maintenance Period as applicable.

Additionally, if a participant has taken a new analgesic/increased regimen of analgesic on 1 or more days (need not be consecutive) prior to the Week 16 visit, then for that week the participant will be classified as an analgesic user. This classification will be used to adjust the formal analysis of the Worst Skin Pain secondary endpoints. If there is a visit date but

no HSSDD available at the visit, then the analgesic/antihistamine user status for that week will be derived based on the visit date. If there is no visit available, then the weekly analgesic/antihistamine user status will default to the analgesic/antihistamine status for the overall study period.

Change #20

Section 8.2 Primary efficacy endpoint

The intercurrent event strategy for the sensitivity analysis was updated in Table 8-1:

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Statistical Category (Section)	Estimands for Primary Endpoint			
	Variable/Endpoint	Pop	IES	PLS (Analysis)
Primary Objective: To evaluate the efficacy of bimekizumab in study participants with moderate to severe HS				
Sensitivity (Section 8.2.3.2)	HiSCR ₅₀ response at Week 16	RS	Composite strategy, as for the primary analysis. The intercurrent events will be handled using a hypothetical strategy , whereby all data at and after the intercurrent event will be treated as missing.	The odds ratio versus placebo based on a logistic regression. Missing values will be imputed using MI – Reference-Based Regression under a missing not at random assumption.

Table 8–1: Estimand Details and Attributes for Primary Endpoint

Statistical Category (Section)	Estimands for Primary Endpoint			
	Variable/Endpoint	Pop	IES	PLS (Analysis)
Sensitivity (Section 8.2.3.5)	HiSCR ₅₀ response at Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants still on randomized treatment at Week 16 are included. whereby only participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

Change #21

Section 8.2.3.5 Analysis on observed cases

The following text was updated:

An additional supportive analysis will be based on observed data only for study participants who are still on the randomized treatment at Week 16. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing (see Section 4.2.2). **participants with a lesion count assessment at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data at Week 16 will be treated as missing (see Section 4.2.2).**

Change #22

Section 8.2.3.11 Center-by-Treatment Interaction

The following sentence was added: In order to achieve model convergence, other explanatory variables eg, Baseline Hurley Stage and Baseline antibiotic use may be dropped from the model.

If model convergence is still not achieved, region and a region-by-treatment interaction term will be added to the model instead. Regions are defined in Section 3.7.

Change #23

Section 8.3 Secondary efficacy endpoints

The following sensitivity analyses as well as the new skin pain response secondary analysis were added to Table 8-2:

Table 813–2: Estimand Details and Attributes for Secondary Analyses

		Estimands for Secondary Endpoints			
Objective Clinical Category	Statistical Category (Section)	Variable/Endpoint	Pop	IES	PLS (Analysis)
Secondary Objective: Evaluate the efficacy of bimekizumab on other measures of disease activity in study participants with moderate to severe HS					
DLQI	Secondary - Sensitivity (Section 8.3.3.2)	Change from Baseline in DLQI total score to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic, rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a DLQI total score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the DLQI total score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.

Table 813–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary (Section 8.3.4.2)	Change from Baseline in worst skin pain score, as assessed by “worst skin pain” item in HSSDD to Week 16	RS	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	Difference in mean change from Baseline (LSMD from ANCOVA) to Week 16 in the worst skin pain score for participants receiving bimekizumab versus placebo. Missing values will not be imputed.
HSSDD	Secondary (Section 8.3.5.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis.	The odds ratio versus placebo based on a logistic regression, as for the primary analysis.

Table 813–2: Estimand Details and Attributes for Secondary Analyses

Objective Clinical Category	Statistical Category (Section)	Estimands for Secondary Endpoints			
		Variable/Endpoint	Pop	IES	PLS (Analysis)
HSSDD	Secondary Sensitivity (Section 8.3.5.2.1)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	Composite strategy , as for the primary analysis where the occurrence of an intercurrent event will be handled by evaluating the corresponding participants as nonresponders.	The odds ratio versus placebo based on a logistic regression. Missing values for any other reason will also be imputed as nonresponders.
HSSDD	Secondary Sensitivity (Section 8.3.5.2.2)	Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, defined as an improvement in the weekly worst skin pain score of at least 3 points	RS ^a	The main intercurrent events are receipt of systemic antibiotic rescue medication or discontinuation of study treatment due to an AE or lack of efficacy prior to Week 16. The intercurrent events will be handled using a while on treatment strategy , whereby only participants with a HSSDD worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16 are included.	The odds ratio versus placebo is based on a logistic regression. Missing values will not be imputed.

AE=adverse event; ANCOVA=analysis of covariance; DLQI=Dermatology Life Quality Index; HiSCR=Hidradenitis Suppurativa Clinical Response; HSSDD=Hidradenitis Suppurativa Symptom Daily Diary; IES=intercurrent event(s) strategy; LSMD=Least Squares Mean Difference; MCMC=Markov Chain Monte Carlo; MI=multiple imputation; PLS=Population-level summary; Pop=Population; RS=Randomized Set

^a **Analysis includes all study participants in the RS with a Baseline HSSDD Worst Skin Pain score of 3 or higher.**

Change #24

Section 8.3.3.1 Primary analysis of change from Baseline in DLQI Total Score

Section 8.3.3.1 was separated from Section 8.3.3 due to the added sensitivity analysis for DLQI Total Score. The following text was added:

Change from Baseline in DLQI total score will be presented by treatment group. The analysis model will be based on an ANCOVA with fixed effects of treatment, Hurley Stage at Baseline, Baseline antibiotic use and Baseline value as a covariate. The least square mean (LSM), standard error (SE), and 95% CI for the LSM will be presented by treatment group. For the comparison between placebo and bimekizumab: the difference between the LSM, the associated 97.5% CI for the contrasts, and the corresponding p-value will be presented. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose with a corresponding $Z_{\alpha/2}$ of 1.96. **Estimand and intercurrent event details are specified in Table 8–2.**

Change #25

The following section was added:

Sensitivity analysis of change from Baseline in DLQI Total score at Week 16 (Section 8.3.3.2)

A sensitivity analysis using the same analysis model as in Section 8.3.3.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

Change #26

Section 8.3.4.1 Primary analysis of change from Baseline in skin pain score at Week 16

Section 8.3.4.1 was separated from Section 8.4.3 due to the added sensitivity analysis for HSSDD worst skin pain score and removed ‘A treatment by analgesic use interaction term will also be added to the model and removed if not significant’ from section.

Change #27

The following section was added:

Sensitivity analysis of change from Baseline in skin pain score at Week 16 (Section 8.3.4.2)

A sensitivity analysis using the same analysis model as in Section 8.3.4.1 will be used, based on observed data. Estimand and intercurrent event details are specified in Table 8–2.

Change #28

The following sections were added:

HSSDD skin pain response at Week 16 (Section 8.3.5)

The analysis set for the analyses of the skin pain response will be restricted to those study participants in the RS with a Baseline worst skin pain score of 3 or higher. The weekly scores and Baseline score are derived as specified in Section 8.3.4.

Primary analysis of skin pain response at Week 16 (Section 8.3.5.1)

Skin pain response at Week 16, as assessed by the “worst skin pain” item in the HSSDD, is defined as an improvement in the weekly worst skin pain score of at least 3 points versus Baseline.

The primary analysis will be based on a logistic regression model including a fixed effect for treatment, Hurley stage at Baseline, Baseline antibiotic use, and analgesic use (Section 6.4.2).

The odds ratio versus placebo, p-value (from Wald test), and 97.5% CI will be calculated. If one dose regimen is tested at the 0.05 significance level as determined in Section 4.5, then the confidence interval will be 95% instead of 97.5% for that dose. Missing data will be handled as specified in Section 4.2.1.2. Estimand and intercurrent event details are specified in Table 8–2.

The number and percentage of participants who are pain responders at Week 16 will be summarized by treatment group.

By-participant listings of pain response status will be provided.

Sensitivity analyses of Skin Pain Response at Week 16 (Section 8.3.5.2)

Nonresponse imputation (Section 8.3.5.2.1)

As a sensitivity analysis, any missing data at Week 16 that are not preceded by an intercurrent event (Table 8–2) will be imputed as nonresponse. That is, participants who experience an intercurrent event will be imputed as nonresponder at the timepoint of the event and all subsequent timepoints (including any recorded data after the event), and all missing data will also be imputed as nonresponse.

The same analysis model as Section 8.3.5.1 will then be used on the imputed data set.

Analysis on observed case (Section 8.3.5.2.2)

An additional supportive analysis will be based on observed data only for study participants with a worst skin pain score at Week 16 who have not had an intercurrent event on or before Week 16. Study participants with missing data or who have prematurely discontinued study treatment will be treated as missing (see Section 4.2.2).

The same analysis model as in Section 8.3.5.1 will then be used on the imputed data set.

Change #29

Section 8.4.1.3 HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ response at both Weeks 16 and 48

The following text was added:

The number and percentage of HiSCR₂₅, HiSCR₅₀, HiSCR₇₅, HiSCR₉₀, and HiSCR₁₀₀ responders at both Weeks 16 and 48 will be summarized based on the RS and MS.

Change #30

Section 8.4.12 Hidradenitis Suppurativa Symptom Daily Diary (HSSDD)

The following text was updated:

See Section 8.3.4 for details on HSSDD Baseline and weekly average definitions and derivations.

Percent change from Baseline in HSSDD responses for worst and average skin pain score is defined as

$$\text{Percent change from Baseline} = 100 \times \frac{\text{Post Baseline HSSDD score} - \text{Baseline HSSDD score}}{\text{Baseline HSSDD score}}$$

Change from Baseline in each HSSDD item (worst skin pain, average skin pain, smell or odor, itch at its worst, and amount of drainage or oozing) score will be summarized using descriptive statistics by treatment group and visit, based on weekly averages. The table will display descriptive statistics for the Baseline, followed by descriptive statistics for the change from Baseline for all visits. Percentage change will be summarized for the worst and average skin pain items.

Additionally, change from Baseline in each HSSDD item will be evaluated by treatment group at Week 16 via continuous empirical cumulative distribution function (eCDF) plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Change from Baseline in Worst **Skin** Pain score and Worst Itch score will additionally be summarized by visit and by analgesic and antihistamine use status (Section 6.4.2), respectively.

HSSDD response **based on clinically meaningful change** for the worst skin pain and average skin pain items is defined as at least a **3-point** 30% reduction and at least a 1-point reduction from Baseline in HSSDD among study participants with a score of ≥ 3 at Baseline, based on weekly averages.

The number and percentage of responders based on clinically meaningful change for **each item the worst skin pain item** will be summarized by treatment group and visit.

The number and percentage of participants who were responders **based on clinically meaningful change** at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst **HSSDD skin pain score item**.

HSSDD response for the worst skin pain and average skin pain items is defined as at least a 30% reduction and at least a 1-point reduction from Baseline among study participants with a score of ≥ 3 at Baseline. The number and percentage of responders for each item will be summarized by treatment group and visit.

The number and percentage of participants who were responders based on clinically meaningful change at any timepoint in the Initial Treatment Period will be summarized by treatment group for the worst skin pain and average skin pain items.

The number and percentage of participants that complete the HSSDD will be calculated for each visit by treatment group. A participant will be counted as completing the HSSDD at a visit if the minimum number of daily entries is present to calculate the weekly average (see Section 8.3.4). The percentage will be based on the number of participants in the RS. A participant will be considered a completer at a visit if the weekly average can be calculated for that visit.

Change #31

Section 8.4.13 Hidradenitis Suppurativa Symptom Questionnaire

The following text was updated:

Additionally, change from Baseline in each HSSQ item will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Missing data for the continuous change from Baseline will be handled by using MI via the MCMC and monotone regression method specified in Section 4.2.1.3.

HSSQ response for skin pain item is defined as at least a 30% reduction and at least a 1-point reduction from Baseline in HS Skin Pain score [11-point numeric rating scale]) among study participants with a score of ≥ 3 at Baseline.

Change #32

Section 8.4.15 Hidradenitis Suppurativa Quality of Life

The following text was added:

Additionally, change from Baseline in each HiSQOL subscale will be evaluated by treatment group at Week 16 and at Week 48 via continuous eCDF plots showing the absolute change from Baseline on the horizontal axis and the cumulative percent of participants experiencing that change on the vertical axis.

Change #33

Section 8.5 Additional statistical analyses of other efficacy endpoints

The following analysis was added

- Skin Pain response per HSSDD at Week 12

Change #34

Section 9.3.2 Anti-bimekizumab antibodies

The section was updated as follows:

Anti-bimekizumab antibodies (ADAb) will be measured using a 3-tiered assay approach: screening assay, confirmatory assay, and titration assay. Samples confirmed as positive within the confirmatory assay will be further evaluated for the presence of neutralizing anti-bimekizumab antibodies specific to IL-17AA, IL-17FF or both. Samples will be taken at Baseline, then at study Weeks 4, 8, 12, 16, 20, 24, 36 and 48, and at PEOT and SFU timepoints.

ADAb samples are not analyzed when study participants are on a treatment other than bimekizumab. For study participants who switch from placebo to bimekizumab, samples are analyzed starting at the visit when the switch to bimekizumab occurs (Week 16). The sample at Week 16 will act as the Baseline for that treatment group.

The screening cut point will be used to determine the status of anti-bimekizumab antibodies in the test sample as Positive Screen (PS) or Negative Screen (NS). For samples presenting anti-bimekizumab antibody levels that are PS, a further confirmatory assay will be performed, and the result of which will be reported as either Positive Immunodepletion (PI) or Negative Immunodepletion (NI).

ADAb status for each sample will be derived as follows:

- Sample values that are either NS, or PS and NI and where the bimekizumab concentration is less than the validated ADAb assay drug tolerance limit will be defined as anti-bimekizumab antibody negative.
- Sample values that are either NS, or PS and NI and where the bimekizumab concentration exceeds the validated ADAb assay drug tolerance limit will be defined as inconclusive.
- Sample values that are PS and PI will be defined as ADAb positive (regardless of availability of a titer value)
- Missing or non-evaluable samples will be defined as missing

Positive immunodepletion samples will be titrated, and the ADAb titer (reciprocal dilution factor including minimum required dilution) will be reported. Subsequently, PI samples will also be subject to a neutralizing assay to evaluate the potential of ADAb to neutralize the target binding of bimekizumab (IL-17AA or IL-17FF or both) in vitro.

Cumulative ~~There are different levels of classification for ADAb status will be derived as follows:~~

The ADAb status (positive, negative or missing) will be considered in a cumulative manner at each time point.

A study participant will be counted positive from the first visit at which the study participant achieved a positive ADAb sample result to the end of the treatment period, regardless of any missing/inconclusive or negative ADAb sample result.

If a study participant has only negative ADAb samples or only one missing/inconclusive sample with all other ADAb samples being negative, the study participant will be classified as negative. An exception remains for the Baseline Visit where only one sample could be available. If the sample is missing/inconclusive, then the sample will be classified as being negative for the cumulative ADAb status.

Otherwise, the study participant will be classified in the missing ADAb category.

Overall ADAb status will be derived as follows:

A study participant will be classified as:

- **Positive if the study participant has at least one positive sample up to the time point of interest (regardless of having missing/inconclusive data).**
- **Negative if the study participant has all the samples negative or only one missing/inconclusive sample with negative ADAb samples up to the timepoint of interest.**
- **Missing if the study participant has more than one missing ADAb result (or have more than one inconclusive sample) and all other available ADAb samples are negative up to the time point of interest.**

ADAb categories will be derived definitions are as follows:

- **Pre ADAb negative – treatment-emergent ADAb negative (Category 1):** Includes study participants who are anti-bimekizumab antibody negative at Baseline and anti-bimekizumab antibody negative at all sampling points during the period of interest (one post-Baseline

missing/inconclusive sample is allowed for subjects with pre- anti-bimekizumab antibody negative sample). This group also includes study participants who have a missing or inconclusive sample (either missing or inconclusive or insufficient volume) at Baseline (ie, pre-treatment) with all post-Baseline samples as ADAb negative.

- **Pre ADAb negative – treatment-emergent ADAb positive (Category 2):** Includes study participants who are ADAb negative at Baseline and ADAb positive at any sampling points post-Baseline during the period of interest. This group also includes study participants who have a missing sample (either missing or insufficient volume) at Baseline (ie, pre-treatment) with 1 or more post-Baseline samples as ADAb positive.
- **Pre ADAb positive – treatment-emergent reduced ADAb (Category 3):** Includes study participants who are ADAb positive at Baseline, and ADAb negative at all sampling points post-Baseline during the period of interest.
- **Pre ADAb positive – treatment-emergent unaffected ADAb positive (Category 4):** Includes study participants who are ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with titer values of the same magnitude as Baseline (ie, less than a predefined fold difference from the Baseline titer).
 - For this analysis, this is set at an increase of less than the validated Minimum Significant Ratio (MSR) of 2.07-fold from Baseline.
- **Pre ADAb positive – treatment-emergent ADAb boosted positive (Category 5):** Includes study participants who ADAb positive at Baseline and are ADAb positive at any sampling point post-Baseline during the period of interest with increased titer values compared to Baseline (equal to or greater than a predefined fold difference increase from Baseline titer which will be defined within the validation of the assay).
 - For this analysis, this is set at an increase equal to or greater than the validated MSR of 2.07-fold from Baseline.
 - Note: for any study participant who is ADAb positive at Baseline and ADAb positive at a post-Baseline time point during the period of interest, but for whom titers are not available to determine treatment unaffected or treatment boosted status, the study participant will be considered as treatment boosted, assuming no other samples are available.
- **ADAb Inconclusive (Category 6):** Includes study participants who have an ADAb positive Baseline (pre-treatment) sample and some post-Baseline samples during the period of interest are missing or inconclusive, while other post-Baseline samples are ADAb negative.
- **Total treatment-emergent ADAb positivity (Category 7 [Categories 2 and 5 combined]):** Includes study participants who are pre ADAb negative – treatment-emergent ADAb positive (Category 2) and pre ADAb positive – treatment boosted ADAb positive (Category 5).
- **Total prevalence of pre- ADAb positivity (Category 8 [Categories 3, 4, 5 and 6 combined]):** Study participants that are tested ADAb positive at Baseline.
- **Missing:** Includes study participants who are ADAb negative, missing, or inconclusive at Baseline with some post-Baseline samples as missing or inconclusive, and other samples as ADAb negative.

For purposes of efficacy subgroup analyses based on anti-bimekizumab antibody status, the following categories can also be used:

- **ADAb positive** – This is defined as study participants who are anti-bimekizumab antibody positive on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).
- **ADAb negative** – Study participants for who either:
 - All samples (including Baseline) are ADAb negative and there are no missing or inconclusive samples
 - Only 1 sample is ADAb positive and all other samples (including Baseline) are ADAb negative or missing or inconclusive
 - Only 1 sample is missing or inconclusive and the remaining ADAb samples are negative.
- **ADAb missing** - Defined as study participants who do not fulfil the criteria for one of the 2 groups listed above.

The rationale for requiring at least 2 time points in which ADAb levels are above the specified cut point is to exclude those study participants who have only one occurrence of ADAb levels during the course of treatment. Including such study participants would increase the number of ADAb positive study participants with potentially no impact on efficacy.

In the case that a sample is collected 1 or more days following the scheduled visit date in which the drug was administered, the ADAb results for that sample will be associated with the scheduled visit and summarized accordingly. Such samples will also be considered when ADAb results are summarized over a given study period.

Analysis

Immunogenicity will be assessed through summary tables and figures, and listing of individual results by participants. All analyses will be run on the AMS, unless specified otherwise.

- Summary of ADAb status overall and by each visit separated by treatment group
- Summary of the time-point of the first occurrence of ADAb positivity during the treatment period by treatment group. A plot of the titer by time to first ADAb positivity will be prepared.
- All individual participant-level ADAb results will be listed.
- The number and percentage of participants in each of the 8 ADAb categories during the treatment period by treatment group.
- The prevalence of immunogenicity, separated by treatment group, and defined subcategory, will be reported by visit, defined as (cumulative) proportion of participants having confirmed positive ADAb samples at any visit up to and including that visit. Missing samples will not be included in the denominator.

- The time to achieving treatment-emergent ADA_b positivity, separated by treatment group and defined subcategory, will be analyzed based on Kaplan-Meier methods. **This will be shown only for Categories 2 and 8 above.** Participants will be considered to have an event at the time point at which treatment emergent ADA_b positive is first achieved (taking the MSR into consideration for sub-category 5). Participants classified as treatment-emergent ADA_b negative will be censored at the time of the last available ADA_b result.
- A summary of HiSCR₅₀ responders at Week 16, separated by treatment group, as a function of ADA_b titer will be presented graphically. ~~This will be repeated for HiSCR₇₅ responders.~~
- Individual plots of plasma bimekizumab concentrations/ ADA_b titer both plotted on the Y-axes by visit (x-axis) for the full treatment period (excluding SFU for interim analyses and including SFU for final analyses) will be presented for participants with and without HiSCR₅₀ response at Week 16.
- Spaghetti plots of ADA_b titer (y-axis) by visit (x-axis), separated by treatment group for all ADA_b positive participants, including Baseline positive participants.
- Box plots of ADA_b titer (logscale) by time to first ADA_b positivity by treatment group.

~~For purposes of efficacy subgroup analyses based on anti bimekizumab antibody status, 2 categories will be used:~~

- ~~• ADA_b positive — This is defined as participants who have ADA_b levels above the specified cut point on at least 2 time points while on treatment (ie, excluding Baseline, excluding SFU).~~
- ~~• ADA_b negative — Participants who are not defined as anti bimekizumab positive (as described above) will be defined as ADA_b negative.~~

The groups for defining ADA_b status for safety subgroup analyses are as follows:

- AEs starting before first ADA_b positive result
- AEs starting on or after first ADA_b positive result
- AEs for participants who were always ADA_b negative

This is further explained in Section 10.2.2.

Change #35

The following section was added:

COVID-19 related considerations (Section 10.2.1.1)

To assess the impact of COVID-19 mass vaccination on TEAEs, a sensitivity analysis will present all TEAEs excluding TEAEs assessed as exclusively related to COVID-19 vaccine by the investigator. TEAEs recorded as related to both study medication and COVID-19 vaccination should not be excluded. A complementary table and listing of TEAEs related to COVID-19 vaccine will be presented.

Another sensitivity analysis will present all TEAEs excluding TEAEs with start date on or up to 5 days after date of COVID-19 vaccine. Note that study participants may receive more than one

administration of COVID-19 vaccine. A complementary table and listing of TEAEs with start date on or up to 5 days after date of COVID-19 vaccine will also be presented.

Change #36

Section 10.2.2 AE summaries

The following AE summaries were added:

- Incidence of TEAEs Excluding TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of TEAEs Exclusively Related to COVID-19 Vaccine by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval Censored TEAEs by SOC, HLT, and PT
- Incidence of COVID-19 Vaccine Interval TEAEs by SOC, HLT, and PT

Change #37

Section 10.2.2 AE summaries

The following text was added:

The following table will be presented for the combined Initial and Maintenance Treatment Period. **This summary will include only AEs that occur while a participant is on bimekizumab. Any AEs in the Initial Treatment Period that begin while a participant is on placebo will be excluded.**

- Incidence of TEAEs per 100 subject years by SOC, HLT, and PT and by Time of Onset Relative to Anti-bimekizumab Antibody Status. **This will include columns for the following:**
 - **TEAEs starting before the first ADAb positive result (includes ADAb categories 2 and 5) where TEAEs have occurred before the following events: a) the first positive ADAb result for subjects in category 2 and b) the first post-Baseline boosted ADAb titer result for subjects with titer results and the first post-Baseline positive ADAb result for subjects with positive ADAb at Baseline with no other samples with titer available for subjects in category 5**
 - **TEAEs starting on the same date or after the first ADAb positive result (includes ADAb Categories 2, 3, 4, 5 and 6) where TEAEs have occurred on or after the following events: a) the first positive ADAb results for subjects in categories 2, 3, 4 and 6, and b) the first post-Baseline boosted ADAb titer result for subjects with titer results and the first post-Baseline positive ADAb result for subjects with positive ADAb at Baseline with no other samples with titer available for subjects in category 5**
 - **TEAEs for subjects who are ADAb negative at all timepoints (includes ADAb Category 1)**

The tables with risk differences will also be accompanied by figures (dot plots) which show the incidence of the adverse events and corresponding 95% risk difference confidence intervals. These will be ordered by descending order of risk difference (bimekizumab vs placebo).

Change #38**Section 10.2.3.3 Major adverse cardiac event**

The entire section was updated:

Potential cardiovascular events are adjudicated by the independent Cardiovascular Event Adjudication Committee (CV-CAC) according to the CV-CAC Charter (version 6.0). Adjudicated events are classified by the CV-CAC to one of the event types as defined in [Table 10–1](#). The classification of an event as a Major Adverse Cardiac Event (MACE) is also determined by the CV-CAC. Events which are classified by the CV-CAC as any of the event types identified in the third column of [Table 10–1](#) will be considered an extended MACE. Note that extended MACE is determined programmatically and includes a broader scope definition of MACE.

MACE as determined by the CV-CAC will be presented in a table and listing. Extended MACE will be presented separately in a table and listing.

Another table and listing will present the adjudicated cardiovascular events by type. For each cardiovascular event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be a MACE and/or an extended MACE.

Additionally, a listing of all events identified for potential review by the CV-CAC will be produced. This listing will indicate whether each event was identified by the CV-CAC Chair for full committee review.

Table 13–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
1	Non-Fatal Myocardial Infarction (MI)	Yes
2	Non-Fatal Stroke: hemorrhagic	Yes
3	Non-Fatal Stroke: ischemic	Yes
4	Non-Fatal Stroke: embolic	Yes
5	Non-Fatal Stroke: undeterminable	Yes
6	Hospitalization or ER for Unstable Angina with urgent revascularization	Yes
7	Hospitalization or ER for Unstable Angina without urgent revascularization	No
8	Hospitalization for Heart Failure	Yes
9	Transient Ischemic Attack (TIA)	No
10	Coronary Revascularization Procedures (e.g. percutaneous coronary intervention, coronary artery bypass grafting)	Yes

Table 13–1: Cardiovascular event classifications

Event Type Code	Event Type	Extended MACE
11	Urgent Revascularization Procedures (i.e. due to symptoms of brain ischemia or pending infarction)	Yes
12	Arrhythmia (not associated with ischemia)	No
13	Peripheral Arterial Event	No
14	Venous Thromboembolic Event: pulmonary embolism (PE)	No
15	Venous Thromboembolic Event: deep vein thrombosis (DVT)	No
16	Venous Thromboembolic Event: PE and DVT	No
17	Other CV Event	No
18	Death due to Myocardial Infarction (MI)	Yes
19	Death due to Stroke	Yes
20	Sudden Cardiac Death	Yes
21	Other CV Death (e.g. heart failure, pulmonary embolism, cardiovascular procedure-related)	Yes
22	Cardiovascular: Undetermined Cause of Death (i.e. cause of death unknown)	Yes
23	Non-Cardiovascular Death	No
24	Non-Cardiovascular Event	No
99	Inadequate information to adjudicate	No

CV=Cardiovascular; DVT=Deep Vein Thrombosis; ER=Emergency Room; MACE= Major Adverse Cardiac Event; MI=Myocardial Infarction; PE=Pulmonary Embolism; TIA=Transient Ischemic Attack.

MACE is determined by the adjudication committee and is not identified programmatically based on event type.

Change #39

Section 10.2.3.5 Suicidal Ideation and Behavior

The entire section was updated:

Potential neuropsychiatric events are adjudicated by the independent Neuropsychiatric Adjudication Committee according to the Neuropsychiatric Adjudication Committee (version 8.0). Adjudicated events are classified by the Committee as Suicidal or Non-suicidal.

Adjudicated events are also further classified by the Committee to one of the event types as

defined in [Table 10–2](#). Suicidal Ideation and Behavior (SIB) is defined as events classified by the Committee as Suicidal.

A table and listing will present SIB events.

Another table and listing will present the adjudicated neuropsychiatric events by type. For each neuropsychiatric event type, the individual PTs which fall within each event type will be summarized. This listing will indicate whether each event was determined to be Suicidal or Non-Suicidal. For event type suicidal ideation, the listing will also indicate if intent was present and if the suicidal ideation was clinically significant.

Additionally, a listing of all events identified for potential review by the Committee will be produced. This listing will indicate whether each event was identified by the Neuropsychiatric Event Adjudication Committee Chair for full committee review.

Table 13–2: Neuropsychiatric event classifications

Event Type Code	Event Classification	Event Type
1	Suicidal	Suicidal events/completed suicide
2	Suicidal	Suicide attempt
3	Suicidal	Preparatory acts toward imminent suicidal behavior
4	Suicidal/Non-suicidal ^a	Suicidal ideation
7	Non-suicidal	Nonsuicidal Self-injurious behavior
8	Non-suicidal	Nonsuicidal Other
99	Not applicable	Inadequate information to adjudicate

^a Suicidal ideation event types can be classified by the Neuropsychiatric Adjudication Committee as Suicidal or Non-suicidal depending on whether intent to die was present.

Change #40

Section 10.2.3.6 Inflammatory bowel disease

The entire section was updated:

Selected gastrointestinal events are adjudicated by the independent Inflammatory Bowel Disease (IBD) Adjudication Committee (IBD-CAC) according to the IBD-CAC Charter (version 3.0). Adjudicated events are classified by the IBD-CAC into one of the diagnostic types as defined in [Table 10–3](#). The events will further be classified as definite, probable or possible IBD.

An overview of adjudicated IBD events will be stratified by subjects with or without a previous medical history of IBD. Previous medical history of IBD will be determined using the information recorded on the History of IBD CRF page ("Does subject have a history of IBD?"). This overview table will present events adjudicated by the IBD-CAC as either possible, probable or definite IBD. Definite and probable IBD will also be aggregated and summarized in this table.

In addition, this table will summarize each IBD event classification (possible, probable or definite) separately.

Another table and listing will present the adjudicated IBD events by type. For each IBD event type, the individual PTs which fall within each event type will be summarized.

Additionally, a listing of all events identified for potential review by the IBD-CAC will be produced. This listing will indicate whether each event was identified by the IBD-CAC Chair for full committee review.

A further supportive listing will present the individual diagnostic criteria met for each adjudicated IBD event.

Table 13–3: IBD event classifications

Event Type Code	Event Type (Classification and diagnosis)	Classification
1	Possible Inflammatory Bowel Disease – Crohn’s Disease	Possible
2	Probable Inflammatory Bowel Disease – Crohn’s Disease	Probable
3	Definite Inflammatory Bowel Disease – Crohn’s Disease	Definite
4	Possible Inflammatory Bowel Disease – Ulcerative Colitis	Possible
5	Probable Inflammatory Bowel Disease – Ulcerative Colitis	Probable
6	Definite Inflammatory Bowel Disease – Ulcerative Colitis	Definite
7	Possible Inflammatory Bowel Disease – type unclassified	Possible
8	Probable Inflammatory Bowel Disease – type unclassified	Probable
9	Definite Inflammatory Bowel Disease – type unclassified	Definite
10	Symptoms not consistent with Inflammatory Bowel Disease	Not applicable
11	Possible Inflammatory Bowel Disease – Microscopic Colitis	Possible
12	Probable Inflammatory Bowel Disease – Microscopic Colitis	Probable
13	Definite Inflammatory Bowel Disease – Microscopic Colitis	Definite
14	Possible Inflammatory Bowel Disease – no further differentiation possible	Possible
15	Probable Inflammatory Bowel Disease – no further differentiation possible	Probable
16	Definite Inflammatory Bowel Disease – no further differentiation possible	Definite
99	Not enough information to adjudicate	Not applicable

IBD=inflammatory bowel disease.

Note: IBD diagnoses of “microscopic colitis” and “no further differentiation possible” were added in an adjudication charter amendment, accounting for the event type numbering.

Change #41

Section 10.2.3.8 Hepatic events and PDILI

The following word was added:

Cases of **potential** Hy's Law will be reported separately in a liver function test table.

Change #42

Section 12.1 Appendix A: MedDRA algorithmic approach to anaphylaxis

The following text was added:

- An **algorithmic approach** which combines a number of anaphylactic reaction symptoms in order to increase specificity. A case must include one of the following where both occur on either the same day as when an injection was administered or one day after, and for scenarios where two events must have been reported, both events must have occurred within one day of each other (**as anaphylaxis is an acute event, imputed dates should not be used in the algorithmic approach**):

STATISTICAL ANALYSIS PLAN SIGNATURE PAGE

This document has been reviewed and approved per the Review and Approval of Clinical Documents Standard Operating Procedures. Signatures indicate that the final version of the Statistical Analysis Plan (SAP) or amended SAP is released for execution.

PUBLIC COPY

This document cannot be used to support any marketing authorization application and any extensions or variations thereof.

Approval Signatures

Name: hs0004-sap-amend-2

Version: 2. 0

Document Number: CLIN-000188631

Title: HS0004 SAP Amendment 2

Approved Date: 07 Jul 2022

Document Approvals	
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 07-Jul-2022 13:31:05 GMT+0000
Approval Verdict: Approved	Name: [REDACTED] Capacity: Clinical Date of Signature: 07-Jul-2022 14:42:54 GMT+0000