

Modification of the EORTC QLQ-C30 (Version 2.0) Based on Content Validity and Reliability Testing in Large Samples of Patients with Cancer

Author(s): D. Osoba, N. Aaronson, B. Zee, M. Sprangers and A. te Velde

Source: *Quality of Life Research*, Mar., 1997, Vol. 6, No. 2 (Mar., 1997), pp. 103-108

Published by: Springer

Stable URL: <https://www.jstor.org/stable/4035330>

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/4035330?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/4035330?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



Springer Research is collaborating with JSTOR to digitize, preserve and extend access to *Quality of Life*

JSTOR

# Modification of the EORTC QLQ-C30 (version 2.0) based on content validity and reliability testing in large samples of patients with cancer

D. Osoba,\* N. Aaronson, B. Zee, M. Sprangers and  
A. te Velde for the Study Group on Quality of Life of the  
EORTC and the Symptom Control and Quality of Life  
Committees of the NCI of Canada Clinical Trials Group

Quality of Life Program, British Columbia Cancer Agency and the University of British Columbia, Vancouver, Canada (D. Osoba); Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands (N. Aaronson, A. te Velde); Clinical Trials Group, National Cancer Institute of Canada and Queens University, Kingston, Canada (B. Zee); Department of Medical Psychology, University of Amsterdam, Amsterdam, The Netherlands (M. Sprangers)

**A revision of the Quality of Life Questionnaire (QLQ-C30) of the European Organization for Research and Treatment of Cancer (EORTC) was undertaken to improve low internal consistency estimates (Cronbach's  $\alpha$ ) and content validity for the role functioning scale and a conceptual difficulty (undue emphasis on physical functioning) in the global quality of life (QOL) scale. The role functioning items were reworded and a four-category response format was substituted for the previous dichotomous format. A new item asking about 'overall health' was substituted for the 'overall physical condition' item in the global QOL domain. The original and new versions were tested at three time points in a total of 1,181 patients with cancer in Canada ( $n = 696$ ) and the Netherlands ( $n = 485$ ). In both samples there was a marked improvement in internal consistency for the role functioning scale (Cronbach's  $\alpha$ s ranging from 0.78–0.88) in the new version. In the global QOL scale, the substitution of the new item for the previous one did not alter internal consistency (Cronbach's  $\alpha$ 's ranging from 0.81–0.92). The revised versions of the role functioning and global QOL domains have been incorporated into the QLQ-C30 (version 2.0).**

**Key words:** Quality of life, QLQ-C30, revisions.

## Introduction

The 30-item version of the core Quality of Life Questionnaire (QLQ-C30) was developed from an earlier version<sup>1</sup> by the Study Group on Quality of Life of the European Organization for Research and Treatment of Cancer (EORTC). Its psychometric properties were tested, initially, by the Study Group in 346 patients with lung cancer<sup>2</sup> and subsequently by the National Cancer Institute of Canada (NCIC) Clinical Trials Group (CTG) in 535 patients with heterogeneous diagnoses (143 with breast cancer, 160 with lung cancer, 111 with ovarian cancer and 121 with other cancers),<sup>3</sup> as well as in Norwegian patients with head and neck cancer<sup>4</sup> and other cancers.<sup>5–10</sup> There were concerns both from the conceptual viewpoint and from the empirical data of the reliability estimates (Cronbach's  $\alpha$ <sup>11</sup>) about the adequacy of the role functioning scale ( $\alpha$  ranging from 0.52–0.66).<sup>2,3</sup> The two items pertaining to role functioning are: 'Are you limited in any way in doing your work or household jobs' and 'Are you completely unable to work at a job or do household jobs?' The response options are 'No' and 'Yes' for both items. Thus, the QLQ-C30 role functioning scale provides a narrow range of possible scores and does not include any questions about the subject's ability to pursue hobbies and leisure time activities. Also, there was a concern, from the conceptual point of view, that one item in the global quality of life scale, asking for a rating of 'overall physical condition', places undue emphasis on physical functioning as a component of this domain, and that it would be desirable to include a

\* To whom correspondence should be addressed at Director, Quality of Life Program, BC Cancer Agency, 600 West 10th Avenue, Vancouver, BC, V5Z4E6, Canada. Tel: (+1) 604 877 6000; Fax: (+1) 604 872 4596; email: dosoba@bccancer.bc.ca

question on overall health in this scale domain.

Therefore, the objectives of the current study were to: (1) revise the wording of the two items in the role functioning scale; (2) restructure the response format of the role functioning scale to a four-category response format; (3) replace the item dealing with 'overall physical condition' in the global quality of life scale with a new item asking about 'overall health'; (4) assess the internal consistency and test-retest reliability of the new items and format and (5) to assess the impact of different questionnaire administration procedures.

## Patients and methods

### Study sample

The study sample consisted of two groups of patients with cancer, one residing in Canada and the other in the Netherlands. The Canadian sample was enrolled in a multicentre study of a 5-hydroxytryptamine (5-HT<sub>3</sub>) receptor antagonist antiemetic and dexamethasone in chemotherapy-naïve patients as part of the first cycle of chemotherapy.<sup>12-13</sup> Patients had to be at least 18 years of age and have an ECOG performance status of 0, 1 or 2. The Dutch sample was recruited specifically for purposes of testing the psychometric properties of the QLQ-C30, as well as several other questionnaires (the Medical Outcomes Survey Short Form-36, the Cancer Rehabilitation and Evaluation System Short Form, and diagnosis-specific questionnaire modules). The patients were beginning either radiotherapy or chemotherapy at the Antoni van Leeuwenhoek Hospital, a specialized cancer treatment centre located in Amsterdam. In the Dutch sample, patients were excluded who: (1) had a life expectancy of less than four months; (2) were considered to be too ill to participate in such a study; (3) lacked basic proficiency in Dutch or (4) were participating in concurrent quality of life investigations. There were no restrictions with regard to age or performance status. The questionnaire was administered either during chemotherapy, immediately after chemotherapy or, when patients were receiving chemotherapy over a period of days, on the last day of chemotherapy. The timing was chosen to coincide with scheduled clinic visits and, when possible, with maximum side effects.

### Revisions to the QLQ-C30

The original wording of the role functioning scale

items was: 'Are you limited in any way in doing either your work or doing household jobs?' and 'Are you completely unable to work at a job or do household jobs?' with the response options being 'Yes' and 'No'. The new items were: 'Were you limited in doing either your work or other daily activities?' and 'Were you limited in pursuing your hobbies or other leisure time activities?' These items were now preceded with the prefix 'During the past week'. The original response format was replaced with 'Not at all', 'A little', 'Quite a bit' and 'Very much' for the new items. In the global quality of life scale, the two original items ('How would you rate your overall physical condition during the past week?' and 'How would you rate your overall quality of life during the past week?') were retained, and a third item, 'How would you rate your *overall health* during the past week?' was added. The seven-category response format for these items remained the same as in the original version. The above items were incorporated into the original version of the QLQ-C30 and this 33-item test version of the questionnaire was labelled the QLQ-C30 (+3).

### Study procedures

Patients in both samples completed the QLQ-C30 (+3) at baseline (Time 1 or T1) within one week before receiving treatment. The questionnaire was readministered in both samples. In Canada, patients were randomized to complete the follow up assessment either on day 4 (T2) or on day 8 (T3) after the first cycle of chemotherapy. In the Netherlands, T2 was approximately 4 weeks after T1, for all patients depending on the treatment schedule (*i.e.*, at the end of radiotherapy or approximately 1 month into the treatment period with chemotherapy). The second follow-up assessment (T3) was carried out 3 months after T2. These assessment points were chosen to capture maximal treatment side effects within a minimal time frame. To assess the impact of different questionnaire administration procedures on the performance of the QLQ-C30, the Netherlands sample was randomly divided into three different modes of administration conditions at T3. One-third of the patients completed the questionnaire during their scheduled visit to the hospital, one-third was administered the questionnaire in the form of a telephone interview, and one-third completed the questionnaire by mail. A subsample of patients completed the QLQ-C30 a fourth time (T4), one week later, for purposes of test-retest reliability estimation. For practical reasons the test-retest sample was

limited to those patients in the mail and telephone conditions. Administrative procedures were held constant across T3 and T4. Research assistants were present during assessments in the clinic or hospital to provide assistance, if necessary, and to record the problems patients may have had with completing the questionnaire.

The languages of the questionnaire were English or Canadian French in Canada and Dutch in the Netherlands.

## Scoring and statistics

The responses to the questionnaire were scored using standard, previously-described, scoring methods.<sup>2</sup> Raw scores were linearly transformed to give values between 0–100. For scores related to functioning domains and global quality of life, higher scores indicate better functioning.

Internal consistency (reliability) of the revised scales and the original scales was determined by Cronbach's coefficient  $\alpha$ .<sup>11</sup> Coefficients higher than 0.7 were considered sufficient for group comparisons.<sup>14</sup> Test–retest reliability was calculated by the intraclass correlation coefficient.<sup>15</sup>

## Results

### Patient characteristics and compliance

The Canadian sample consisted of 696 patients with heterogeneous cancer diagnoses in a multicentre study of a new 5-HT<sub>3</sub> antiemetic. All patients agreed to participate and completed the baseline assessment; enrolment took place between 5 May 1993 and 25 January 1995. All patients completed the assessment at Time 1 (T1), 336 patients (96.6%) completed the assessment at T2 and 316 completed it at T3 (90.9%).

The Netherlands Cancer Institute sample included 485 patients out of 614 who were approached (79%) for enrolment between September 1992 and April 1994. Reasons for declining study participation included: (1) the study was perceived as too emotionally burdensome ( $n = 54$ ); (2) perceived lack of time ( $n = 22$ ); (3) lack of interest ( $n = 18$ ) or (4) being too ill ( $n = 10$ ). The remaining 25 patients had a variety of other reasons. Patients declining participation were, on average, older (mean age 65 years vs. 57 years), were less frequently married (59% vs. 76%), and more often had compulsory education only (91% vs. 82%) than those who participated.

Of the 485 patients enrolled in the Dutch sample, 445 (92%) completed the second assessment, and 369 (76%) completed the third assessment. Of the patients answering the questionnaire at T3, 92 were expected to complete the questionnaire at T4. Of these 83 did so for a completion rate of 93.5%. The primary reasons for patient attrition in the Dutch sample were severe illness or death. Patients lost to follow-up were more likely to have metastatic disease, and their KPS scores were 10–30 points lower than in patients who continued participation. It is unlikely that potential differences between the original and the new scales would be affected by drop out rates. The average time between T1 and T2 was 30 days (SD = 9 days), between T2 and T3 was 98 days (SD = 13 days), and between T3 and T4 was 7 days (SD = 5 days).

The characteristics of both samples are presented in Table 1. There was good representation across age groups, primary tumour sites and extent of disease.

### Scale levels

The change in response format for the role functioning scale from a dichotomous format in the QLQ-C30 to a four-category response format in the QLQ-C30 (+3) increased the range of scores within the role functioning scale (*i.e.*, from three scale levels for the dichotomous response option to seven scale levels with the four-category response option). Since the revision to the global QOL scale involved only the replacement of one item with another, the number of scale levels remained the same (*i.e.*, 13). There was no increase in the number of items for the role of the global QOL scales.

### Internal consistency

*Role functioning.* The full range of responses was observed (0–100) for each of the two items in both the original and new role functioning scales. The mean scores at T1 were similar in both patient samples, but at T2 were lower in the Canadian sample than in the Netherlands sample. However, at any one time point, the mean scores for the original and new versions in each country were similar, while the standard deviation was consistently smaller in the new version (Table 2). This likely reflects the increase in the number of scale response levels (from three to seven) between the original and new role functioning scales.

The change from the original items and dichotomous response format for the role functioning scale to the revised items and four-category response

format resulted in a dramatic increase in Cronbach’s  $\alpha$  at all three time points in both cohorts. Cronbach’s  $\alpha$  estimates were between 0.59–0.67 in the Canadian cohort and between 0.26–0.35 in the Dutch cohort for the original items, but increased to between 0.78–0.88 in both cohorts in the new version (Table 2).

*Global QOL.* A comparison of the original global QOL scale with the revised scale showed responses ranging from 0–100 for each item with similar means and standard deviations (Table 2) for both versions in both countries. Cronbach’s  $\alpha$  were similar for the original and revised scales at all three time points,

Table 1. Patient characteristics

	Canada (n = 696)		The Netherlands (n = 485)	
Age				
Mean years	55	—	57	—
(Range)	(18–86)	—	(22–86)	—
Karnofsky status				
Mean score	—	—	78	—
(Range)	—	—	(30–100)	—
	No.	%	No.	%
Gender				
Male	205	29.5	205	42.3
Female	491	70.5	280	57.7
Primary site				
Breast	299	43.0	170	35.1
Lung	197	28.3	150	30.9
Colorectal	0	0.0	117	24.1
Gynaecologic	76	10.9	0	0.0
Other	124	17.8	48	9.9
Extent of disease				
Local/regional	337	48.4	296	61.0
Distant metastases	358	51.3	189	39.0
Unknown	1	0.2	0	0.0
ECOG performance status				
0	346	49.7	—	—
1	281	40.4	—	—
2	69	9.9	—	—
Treatment				
Radiation therapy	0	0.0	242	49.9
Chemotherapy	696	100.0	243	50.1

\* Percentages may not add up to 100 due to rounding

Table 2. Internal consistency (reliability) of the original QLQ-C30 and the new items/response formats of the QLQ-C30 (+3)

Scale	Time	Canada (n = 696)						The Netherlands (n = 485)					
		C30			C30 (+3)			C30			C30 (+3)		
		Mean	(SD)	$\alpha$	Mean	(SD)	$\alpha$	Mean	(SD)	$\alpha$	Mean	(SD)	$\alpha$
Role functioning	1	70.2	(32.5)	0.61	70.4	(30.1)	0.82	66.7	(32.5)	0.26	64.8	(30.1)	0.78
	2	52.8	(42.3)	0.67	57.8	(33.1)	0.85	64.9	(33.3)	0.35	62.8	(30.7)	0.83
	3	67.9	(37.9)	0.59	68.6	(31.8)	0.88	66.5	(33.0)	0.29	69.5	(30.6)	0.86
Global QOL (2-item)	1	61.7	(23.0)	0.83	62.9	(23.0)	0.83	63.9	(21.1)	0.81	63.5	(21.6)	0.79
	2	52.9	(24.0)	0.88	53.4	(24.0)	0.90	62.1	(21.5)	0.86	62.5	(21.5)	0.84
	3	59.9	(23.6)	0.90	60.4	(23.8)	0.92	66.6	(21.9)	0.86	67.0	(22.4)	0.87



ranging from 0.83–0.90 for the C30 and 0.83–0.92 for the C30 (+3) in the Canadian sample and from 0.81–0.86 for the C30 and 0.79–0.87 for the C30 (+3) in the Dutch sample.

*Response stability.* The relevant scales of the C30 and the C30 (+3) were also tested for response stability (test–retest reliability) over time in a subpopulation of 86 patients from the Dutch sample who were known to be in a stable condition (no treatment and no change in disease status). The mean interval between test and retest was 7 days (SD = 5 days). The intraclass correlation coefficients for the role functioning scale were 0.59 for the C30 and 0.82 for the C30 (+3). For the global QOL scale they were 0.82 for the C30 and 0.83 for the C30 (+3). Only very minor differences were noted in test–retest reliability estimates as a function of mode of administration (data not shown).

## Discussion

Initial versions of health-related quality of life (HRQOL) questionnaires may require revision after some experience in use. This is viewed as a normal process during questionnaire development. The first version of the EORTC QLQ-C30 was derived from a previous 36-item version.<sup>1</sup> After trials in large numbers of cancer patients with a variety of diagnoses and stages of disease, it was evident that the internal consistency (Cronbach's  $\alpha$ ) of the two-item role functioning scale in the QLQ-C30 was inadequate<sup>2–3</sup> and that revision was required. Thus, the two items were reworded and, in addition, the response format was changed from a dichotomous to a four-category format, identical to the one used in most of the remainder of the questionnaire. The results from tests of the revised role functioning scale indicate a much higher Cronbach's  $\alpha$ . Thus, in the new version (2.0) of the QLQ-C30, the previous role functioning items have been deleted and the revised items substituted in their place. The mean scores for the revised scale were very similar to those of the original version, as assessed at three time points in two cohorts of patients (Canadian and Dutch). Due to the greater number of possible response levels with the four-category response format, the standard deviations of the revised scales were smaller. Despite there being no increase in the number of items in the revised scale, these results clearly are an improvement over those obtained from the original role functioning scale.

The reason for revision of the global QOL scale

was the undue emphasis on physical condition placed by one of the items rather than a low internal consistency. Cronbach's  $\alpha$  for the original version in previous studies has been near 0.80 or higher.<sup>2–6</sup> In order to not emphasize the contribution of physical condition to global QOL, this item was replaced by one asking about 'overall health' while the other item asking about 'overall quality of life' was retained. The substitution of the revised item for the original one did not alter the group mean scores, standard deviations or Cronbach's  $\alpha$ s in either cohort at each of the three time points. Thus, in version 2.0 of the QLQ-C30, the new item has been substituted for the original one.

Testing of the original QLQ-C30 (+3) for response stability over time (test–retest reliability) in a subpopulation of the Netherlands sample, gave a higher intraclass correlation coefficient (0.82) for the new role functioning scale than for the original one. Intraclass correlations for the original and revised versions of the global QOL scale were both high. These results were similar to those reported for a Nordic sample.<sup>7</sup> Thus, these scales of the QLQ-C30 (version 2.0) appear to have good test–retest reliability.

The revisions to the role functioning and global QOL scales described in this paper are incorporated into the QLQ-C30, version 2.0.<sup>16</sup> For those investigators who have previously used or are currently using the QLQ-C30 (version 1.0) in their research, and who may wish to compare results from these studies with those of future studies, the EORTC will continue to make available the interim version of the questionnaire — the QLQ-C30 (+3) — which contains both the original and revised role functioning and global QOL scales. Where such comparisons across studies are not planned, we would recommend employing the QLQ-C30 (version 2.0) in which the original role functioning and global QOL scales have been replaced with the revised versions of these scales.

## References

1. Aaronson NK, Ahmedzai S, Bullinger M, *et al.* The EORTC core quality of life questionnaire: interim results of an international field study. In: Osoba D, ed. *Effect of Cancer on Quality of Life*. Boca Raton, FL (USA): CRC Press, 1991: 185–203.
2. Aaronson NK, Ahmedzai S, Bergman B, *et al.* The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international trials in oncology. *J Natl Cancer Inst* 1993; **85**: 365–376.
3. Osoba D, Zee B, Pater J, *et al.* Psychometric properties and responsiveness of the EORTC Quality of Life Questionnaire (QLQ-C30) in patients with breast, ovarian and lung cancer. *Qual Life Res* 1994; **3**: 353–364.

4. Bjordal K, Kaasa S. Psychometric validation of the EORTC Core Quality of Life Questionnaire, 30-item version and a diagnosis-specific module for head and neck cancer patients. *Acta Oncol* 1992; **31**: 311–321.
5. Ringdal GI, Ringdal K. Testing the EORTC Quality of Life Questionnaire on cancer patients with heterogenous diagnoses. *Qual Life Res* 1993; **2**: 129–140.
6. Niezgodna HE, Pater JL. A validation study of the domains of the core EORTC Quality of Life questionnaire. *Qual Life Res* 1993; **2**: 319–325.
7. Hjermstad MJ, Fossa SD, Bjordal K, et al. Test/retest study of the European Organization for Research and Treatment of Cancer core Quality-of-Life Questionnaire. *J Clin Oncol* 1995; **13**: 1249–1254.
8. Schaafsma J, Osoba D. The Karnofsky performance status scale re-examined: a cross-validation with the EORTC-C30. *Qual Life Res* 1994; **3**: 413–424.
9. Kaasa S, Bjordal K, Aaronson N, et al. The EORTC Core Quality of Life Questionnaire (QLQ-C30): Validity and reliability when analyzed with patients treated with palliative radiotherapy. *Eur J Cancer* 1995; **31A**: 2260–2263.
10. Groenvold M, Bjorner JB, Klee MC, Kreiner S. Test for item bias in a quality of life questionnaire. *J Clin Epidemiol* 1995; **48**: 805–816.
11. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; **16**: 297–334.
12. Lofters W, Zee B. Dolasetron vs. ondansetron with or without dexamethasone in the prevention of nausea and vomiting in patients receiving moderately emetogenic chemotherapy [abstract]. *Support Care Cancer* 1995; **3**: 338.
13. Osoba D, Pater JL, Zee B, et al. Effective antiemetic therapy improves quality of life (QOL) after moderately emetogenic chemotherapy (MEC) [abstract]. *Qual Life Res* 1995; **4**: 467–468.
14. Nunally JC, Bernstein IH. *Psychometric Theory*. Third Edition. New York, NY (USA): McGraw-Hill, 1994: 264–265.
15. Steel RGD, Torrie JH. *Principles and Procedures of Statistics*. Second Edition. New York, NY (USA): McGraw-Hill, 1960.
16. Fayes P, Aaronson N, Bjordal K, Sullivan M. *EORTC QLQ-C30 Scoring Manual*. Brussels, Belgium: EORTC Study Group on Quality of Life, 1995.

(Received 29 September 1996;  
accepted 2 January 1997)