

MINISTÉRIO DA SAÚDE

**DIRETRIZES METODOLÓGICAS: Elaboração de revisão sistemática e metanálise
de estudos observacionais comparativos sobre fatores de risco e prognóstico**

Brasília – DF
2014



MINISTÉRIO DA SAÚDE
Secretaria de Ciência, Tecnologia e Insumos Estratégicos
Departamento de Ciência e Tecnologia

DIRETRIZES METODOLÓGICAS

**Elaboração de revisão sistemática e metanálise de
estudos observacionais comparativos sobre fatores
de risco e prognóstico**

Brasília – DF
2014





Esta obra é disponibilizada nos termos da Licença Creative Commons – Atribuição – Não Comercial – Compartilhamento pela mesma licença 4.0 Internacional. É permitida a reprodução parcial ou total desta obra, desde que citada a fonte.

A coleção institucional do Ministério da Saúde pode ser acessada, na íntegra, na Biblioteca Virtual em Saúde do Ministério da Saúde: <www.saude.gov.br/bvs>.

Este trabalho foi desenvolvido no âmbito do termo de cooperação nº 47 entre o Departamento de Ciência e Tecnologia e a Organização Panamericana da Saúde

Tiragem: 1ª edição – 2014 – 1.000 exemplares

Elaboração, distribuição e informações:

MINISTÉRIO DA SAÚDE

Secretaria de Ciência, Tecnologia e Insumos Estratégicos

Departamento Ciência e Tecnologia

Coordenação-Geral de Gestão do Conhecimento

SQN Quadra 2 Projeto C, térreo sala 04

CEP: 70712-902 – Brasília/DF

Tel: (61) 3410-4199

Site: www.saude.gov.br

E-mail: ats.decit@saude.gov.br

Supervisão Geral:

Carlos Augusto Gabrois Gadelha (SCTIE/MS)

Antônio Carlos Campos de Carvalho (Decit/SCTIE/MS)

Jorge Otávio Maia Barreto (Decit/SCTIE/MS)

Organização:

Kathiaja Miranda Souza (Decit/SCTIE/MS)

Roberta Moreira Wichmann (Decit/SCTIE/MS)

Elaboração:

Érica Aranha Suzumura (IEP-HCor)

Anna Maria Buehler (IEP-HCor)

Sabrina Bernardes Pereira (IEP-HCor)

Mabel Figueiró (IEP-HCor)

Claudia Oliveira (IEP-HCor)

Alexandre Biasi Cavalcanti (IEP-HCor)

Otávio Berwanger (IEP-HCor)

Revisão Técnica:

Adriano Alves Pereira (Universidade Federal de Uberlândia)

Airton Stein (NATS/GHC; UFCSPA; ULBRA)

Marina Gonçalves de Freitas (Decit/SCTIE/MS)

Mário Henrique Osanai (Decit/SCTIE/MS)

Nashira Campos Vieira (ANVISA)

Roberta Moreira Wichmann (Decit/SCTIE/MS)

Revisão de Especialista:

Rachel Riera (Cochrane/Brasil)

Editoração:

Eliana Carlan (Decit/SCTIE/MS)

Jessica Alves Rippel (Decit/SCTIE/MS)

Design Gráfico:

Gustavo Veiga e Lins (Decit/SCTIE/MS)

Normalização:

Amanda Soares Moreira (CGDI/ Editora MS)

Impresso no Brasil/Printed in Brazil

Ficha Catalográfica

Brasil. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia.

Diretrizes metodológicas : elaboração de revisão sistemática e metanálise de estudos observacionais comparativos sobre fatores de risco e prognóstico / Ministério da Saúde, Secretaria de Ciência, Tecnologia e Insumos Estratégicos, Departamento de Ciência e Tecnologia. – Brasília : Ministério da Saúde, 2014.

132 p. : il.

ISBN 978-85-334-2171-4

1. Avaliação de tecnologias em saúde. 2. Diretrizes metodológicas. 3. Tecnologias em saúde. I. Título.

CDU 001.8:614

Catalogação na fonte – Coordenação-Geral de Documentação e Informação – Editora MS – OS 2014/0529

Títulos para indexação

Em inglês: Methodological guideline: how to conduct a systematic review and meta-analysis of observational comparative studies of risk factors and prognosis

Em espanhol: Directriz metodológica: cómo preparar un revisión sistemática y meta-análisis de estudios observacionales comparativos de factores de riesgo y pronóstico

Lista de siglas e abreviaturas

AMI – Australasian Medical Index

BVS – Biblioteca Virtual em Saúde

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CMA – Comprehensive Meta-Analysis

CBLD – Chinese Biomedical Literature Database

CRD – Centre for Reviews and Dissemination

CINHAL – Cumulative Index to Nursing and Allied Health Literature

DeCS – Descritores em Ciências da Saúde

EMBASE – Base de dados eletrônica da editora Elsevier

Emtree – Descritores de assunto do Embase

GRADE – Grading the quality of evidence and the strength of recommendations

HIRU – Health Information Research Unit

IBICT – Instituto Brasileiro de Ciência e Tecnologia

ISRCTN – International Standard Randomised Controlled Trial Number Register

JAMA – The Journal of the American Medical Association

LILACS – Latin American and Caribbean Health Sciences Literature

MEDLINE – Medical Literature Analysis and Retrieval System Online

MeSH – Medical Subject Headings

MOOSE – Meta-analysis of observational studies in epidemiology

Open SIGLE EAGLE – European Association for Grey Literature Exploitation

OVID – Plataforma de acesso para base de dados disponibilizada via assinatura da editora Wolters Kluwer Health

PLOS Medicine – Public Library Of Science Medicine

PubMed – Serviço de Biblioteca Nacional de Medicina dos Estados Unidos para acesso gratuito ao Medline

PROSPERO – International Prospective Register of Ongoing Systematic Reviews

ReBEC – Registro Brasileiro de Ensaio Clínicos

Rebrats – Rede Brasileira de Avaliação de Tecnologias em Saúde

RevMan – Review Manager

SAS – Statistical Analysis System

SCOPUS – Base de dados de resumos e citações de literatura científica

SciELO – Scientific Electronic Library Online

STROBE – STrengthening the Reporting of OBservational studies in Epidemiology

TRIP – Turning Research into Practice Database

LISTA DE QUADROS

Quadro 1 – Diferenças entre revisões sistemáticas e tradicionais	15
Quadro 2 – Estratégia PECO	21
Quadro 3 – Exemplo de questão de pesquisa de estudos observacionais	22
Quadro 4 – Estrutura de protocolo sugerida pela Cochrane Library	24
Quadro 5 – Características presentes em estudos observacionais	34
Quadro 6 – Variáveis consideradas para elaboração da ficha de extração dos dados	39
Quadro 7 – Principais vieses encontrados em revisões sistemáticas	50
Quadro 8 – Estruturação da discussão da revisão sistemática	60
Quadro 9 – Sistema GRADE para avaliação da qualidade das evidências	63
Quadro 10 – Fatores que diminuem a qualidade da evidência	64
Quadro 11 – Fatores que aumentam a qualidade da evidência	65

LISTA DE FIGURAS

Figura 1 – Exemplo de registro de protocolo no ClinicalTrials.gov	29
Figura 2 – Apresentação esquemática da variável de confusão e sua relação com exposição e com desfecho	38
Figura 3 – Gráfico de funil hipotético	51

SUMÁRIO

Apresentação	13
1 Introdução às revisões sistemáticas	15
1.1 Definição de revisão sistemática	15
1.2 Definição de metanálise	15
1.3 Importância das revisões sistemáticas de estudos observacionais comparativos	16
1.4 Tipos de estudos incluídos em revisões sistemáticas sobre fatores de risco e prognóstico	17
1.5 Limitações das revisões sistemáticas de estudos observacionais	18
1.6 Recursos necessários	19
2 Planejamento da revisão sistemática	21
2.1 Definição da questão de pesquisa estruturada no formato do acrônimo PECO	21
2.2 Definição dos critérios de elegibilidade	22
2.3 Revisão de literatura: justificativa para a revisão sistemática	22
2.4 Documentação da metodologia: redação do protocolo	23
3 Condução da revisão sistemática	27
3.1 Busca de potenciais estudos elegíveis	27
3.1.1 Base de dados bibliográficas eletrônicas	27
3.1.2 Pesquisa em fontes de dados não publicados, em andamento e literatura cinzenta	29
3.1.3 Elaboração da estratégia de busca nas várias bases de dados	30
3.1.3.1 Seleção dos termos para a busca	31
3.1.3.2 Filtros de busca para estudos observacionais	32
3.1.4 Relatando o processo de busca	32
3.2 Avaliação da elegibilidade dos estudos	33
3.2.1 Triagem dos artigos pela leitura de título e resumo	33
3.2.2 Confirmação da elegibilidade: leitura do artigo completo	35
3.3 Extração de dados	36
3.3.1 Processo de extração de dados	36
3.3.1.1 Elaboração da ficha de extração de dados	37
3.3.1.2 Características dos estudos incluídos e síntese qualitativa dos dados	39
3.3.1.3 Avaliação do risco de viés dos estudos incluídos	40
3.3.1.4 Extração dos dados quantitativos	42
3.4 Resultados	43
3.4.1 Apresentação dos dados descritivos	43
3.4.2 Sumário das medidas de associação nos estudos	44
3.4.2.1 Cálculo da metanálise de estudos observacionais	44
3.4.2.2 Modelos de análise	47

3.4.2.2.1 Modelo de efeito fixo	47
3.4.2.2.2 Modelo de efeitos randômicos	47
3.4.2.3 Métodos estatísticos	47
3.4.2.3.1 Método do inverso da variância	47
3.4.2.3.2 Método Mantel-Haenszel	48
3.4.2.3.3 Método de Peto	48
3.4.2.4 Programas para cálculo de metanálise	48
3.4.3 Apresentação dos resultados das metanálises	49
3.4.4 Risco de viés de publicação	49
3.4.4.1 Gráfico de funil	50
3.4.4.2 Teste para assimetria do gráfico de funil	52
3.4.4.3 Método “trim” e “fill”	53
3.4.5 Heterogeneidade	54
3.4.5.1 Quantificação de heterogeneidade	54
3.4.5.2 Explorando a heterogeneidade entre os estudos	55
3.4.5.2.1 Meta-regressão	55
3.4.5.2.2 Análise de subgrupo (ou análise estratificada)	56
3.4.6 Análise de sensibilidade	57
4 Relato e aplicabilidade dos resultados	59
4.1 Estruturando a discussão e a conclusão.....	59
4.1.1 Estruturando a discussão	59
4.1.2 Estruturando a conclusão	60
4.1.2.1 Implicações para prática clínica	60
4.1.2.2 Implicações para pesquisas futuras	61
4.2 Utilização da diretriz de redação do manuscrito – MOOSE	61
5 Etapas opcionais	63
5.1 Avaliação da qualidade da evidência – sistema GRADE	63
5.1.1 O sistema GRADE e a força da recomendação	64
5.1.2 Como graduar a qualidade da evidência	64
Referências	67
Apêndices	75
Apêndice A – Principais tipos de estudos observacionais para avaliação de risco e prognóstico	77
Apêndice B – Exemplo de sensibilização de busca utilizando os “entry terms” do termo MeSH para “ <i>brain natriuretic peptide</i> ”	79
Apêndice C – Construção da estratégia de busca: símbolos e operadores booleanos	81
Apêndice D – Gerenciadores de referências: o que são e como utilizá-los.....	85
Apêndice E – Exemplo de ficha de avaliação de elegibilidade	89
Apêndice F – Cálculo e interpretação das medidas de associação	91
Apêndice G – Interpretando o gráfico de floresta	95

Anexos 97

 Anexo A – Strengthening the Reporting of Observational studies in
 Epidemiology (STROBE) 99

 Anexo B – Fluxo dos artigos da revisão sistemática103

 Anexo C – Newcastle-Ottawa: Escala de avaliação de qualidade105

 Anexo D – Checklist Downs and Black adaptado para avaliação da
 qualidade metodológica111

 Anexo E – Diretriz de redação do manuscrito – Meta-analysis of
 observational studies in epidemiology (MOOSE)²127

As revisões sistemáticas são estudos secundários que pretendem sumarizar as evidências existentes sobre a associação entre uma exposição e desfechos, por meio de metodologia confiável, rigorosa e reprodutível. Tem por objetivo localizar, avaliar criticamente e interpretar todos os estudos disponíveis para uma questão de pesquisa, área do conhecimento ou fenômeno de interesse. O resultado dessa revisão representa o atual conhecimento sobre determinado fator de exposição (síntese qualitativa), enquanto o método estatístico geralmente empregado (metanálise) sumariza em uma única medida de associação os resultados de diferentes estudos (síntese quantitativa), elevando a precisão dessas estimativas.

A Oficina de Prioridades de Pesquisa em Saúde, “Editais Temáticos”, realizada em março de 2006, definiu a importância das diretrizes metodológicas para a elaboração de revisões sistemáticas. O seu principal objetivo foi padronizar e equalizar a qualidade na elaboração e condução desses estudos.

Por meio desta iniciativa, o Departamento de Ciência e Tecnologia (Decit) criou o Grupo de Trabalho de Métodos da Rede Brasileira de Avaliação de Tecnologias em Saúde (Rebrats), que, dentre outras atribuições, ficou responsável pela Elaboração de Diretrizes Metodológicas. O grupo é formado por profissionais de diversas especialidades e instituições, tais como hospitais de excelência, universidades federais, fundações e técnicos de secretarias e departamentos do Ministério da Saúde.

O objetivo deste documento é apresentar os principais conceitos utilizando uma linguagem mais simples e acessível para os profissionais que estejam ou não familiarizados com a metodologia envolvida na condução de revisões sistemáticas e metanálises sobre fatores de risco e prognóstico.

A diretriz apresentada neste documento foi baseada em diretrizes prévias, ferramentas de avaliação de qualidade metodológica e de relato adequado dos resultados¹⁻³. Adicionalmente, foram consultados artigos de revisão metodológica publicados em periódicos científicos. A estrutura desta Diretriz está organizada em cinco seções:

- Seção 1: Introdução às revisões sistemáticas como um método de pesquisa.
- Seção 2: Planejamento de uma revisão sistemática.
- Seção 3: Condução da revisão sistemática e seus resultados.
- Seção 4: Relato e aplicabilidade dos resultados.
- Seção 5: Etapas opcionais complementares.

Essa diretriz visa orientar e padronizar a elaboração e condução de uma revisão sistemática sobre fatores de risco e prognóstico. No entanto, ela não leva em consideração a relevância do tipo de questão no processo da revisão e não esgota todos os mecanismos necessários para entender e realizar uma metanálise. O detalhamento de alguns conceitos fundamentais necessários ao entendimento desse manual pode ser encontrado nos apêndices.

1.1 Definição de revisão sistemática

Revisão sistemática é um tipo de estudo secundário que utiliza um processo abrangente de revisão da literatura, de forma imparcial e reproduzível, para localizar, avaliar criticamente e sintetizar o conjunto de evidências disponível na literatura científica para obter uma visão geral sobre determinada questão de pesquisa. Diferencia-se da revisão narrativa ou revisão tradicional uma vez que inclui uma descrição detalhada dos métodos e critérios utilizados para selecionar e avaliar os artigos incluídos. A revisão narrativa tem como viés a abordagem da perspectiva do autor, uma vez que não tem o compromisso de apresentar de forma sistemática toda a evidência disponível sobre a questão de pesquisa. As principais diferenças entre revisões sistemáticas e revisões narrativas encontram-se no Quadro 1.

Quadro 1 – Diferenças entre revisões sistemáticas e tradicionais

Características	Revisão Tradicional	Revisão Sistemática
Questão clínica	Raramente reportada. Geralmente são amplas, abrangendo vários aspectos sobre determinada população, doença ou exposição.	Questão clínica objetiva, com delimitação da população, exposição e desfecho de interesse.
Busca na literatura	Raramente reportada e não sistematizada.	Busca abrangente em diversas bases bibliográficas.
Seleção de estudos	Raramente reportada. Geralmente com amostra enviesada de estudos.	Crítérios bem definidos de inclusão e exclusão de estudos.
Avaliação da qualidade dos artigos selecionados	Raramente reportada. Não sistematizada.	Avaliação metodológica criteriosa dos artigos selecionados.
Sumário dos resultados	Usualmente qualitativo.	Qualitativa e/ou quantitativa. Se quantitativa é expressa por metanálises.

Fonte: (Guyatt G., 2002, adaptado)⁴.

1.2 Definição de metanálise

Metanálise é uma análise estatística que sumariza as medidas de associação de dois ou mais estudos independentes, gerando uma única medida de associação⁵. No contexto de metanálise de estudos observacionais, é importante destacar que o número desse tipo de estudo tem aumentado substancialmente nas últimas quatro décadas, sendo útil para situações nas quais é necessário sintetizar os resultados de áreas de pesquisa clínica não passíveis de conduzir ensaios clínicos randomizados. No entanto, a escolha do delineamento observacional em pesquisas para inferir causalidade ainda é controversa⁶.

Dada a natureza dos estudos observacionais, a metanálise desses estudos apresenta desafios particulares devido aos vieses inerentes às análises e às diferenças no delineamento e nos resultados dos estudos incluídos (heterogeneidade clínica ou estatística, respectivamente), considerando os diversos fatores de confusão presentes. Obviamente, o aumento do tamanho da amostra é uma meta a fim de obter estimativas mais precisas acerca da associação entre o fator de exposição e o desfecho.

Nesse sentido, uma metanálise de estudos observacionais comparativos tem como propósito principal a investigação das potenciais fontes de heterogeneidade. Assim como as revisões sistemáticas de ensaios clínicos randomizados, a metanálise deve ser considerada apenas quando os estudos são similares, tanto em relação às características clínicas dos pacientes incluídos, quanto em relação à intervenção/exposição em estudo. Mesmo com esses critérios preenchidos, ainda pode haver heterogeneidade importante entre os estudos, tornando inapropriado sumarizar as medidas de associação em uma única estimativa. Nesse caso, os resultados dos vários estudos são apresentados de forma descritiva.

Finalmente, uma revisão sistemática precisa cumprir vários requisitos fundamentais para garantir a validade de seus resultados: a) objetivos bem definidos, incluindo definições precisas de variáveis clínicas e resultados; b) busca e seleção adequada dos estudos, com uso de estratégia de busca sensível e bem documentada; c) avaliação dos vieses na identificação e seleção dos estudos; d) descrição e avaliação de heterogeneidade; e) justificativa do uso dos métodos de avaliação dos dados; f) realização de análises de subgrupo ou de sensibilidade, principalmente em metanálise de estudos observacionais comparativos a fim de identificar ou isolar os potenciais fatores confundidores⁷.

1.3 Importância das revisões sistemáticas de estudos observacionais comparativos

Com o advento da medicina baseada em evidências, têm crescido a necessidade de síntese de evidências e respostas rápidas às questões clínicas antes da sua aplicação na prática. São necessárias evidências confiáveis para embasar a tomada de decisão em saúde, tanto na medicina como nas áreas afins (enfermagem, nutrição, fisioterapia, psicologia, fonoaudiologia, terapia ocupacional, entre outras). Para tanto, revisões sistemáticas e metanálises têm sido frequentemente utilizadas no intuito de identificar, avaliar e sintetizar os resultados de estudos clínicos conduzidos para uma questão de interesse.

Um ensaio clínico randomizado é o delineamento de pesquisa ideal para avaliação dos efeitos de intervenções. No entanto, ensaios clínicos randomizados não são adequados ou viáveis para responder determinadas questões de pesquisa, em particular sobre fatores de risco ou prognóstico. Isso decorre de três principais razões: a) pode ser considerado não ético randomizar pacientes para expô-los a potenciais fatores de riscos; b) alguns desfechos ou doenças podem apenas ocorrer

após a inclusão de grande amostra populacional e longos períodos de seguimento, o que pode inviabilizar um ensaio clínico randomizado; c) dependendo da exposição de interesse, muitas pessoas poderiam se recusar a receber a exposição determinada por terceiros, por longo período de tempo⁸. Assim, nas situações em que não seja factível a realização de ensaios clínicos randomizados (por questões éticas, operacionais ou financeiras), os estudos observacionais apresentam-se como uma alternativa adequada.

1.4 Tipos de estudos incluídos em revisões sistemáticas sobre fatores de risco e prognóstico

Em geral, as revisões sistemáticas de intervenções priorizam os ensaios clínicos randomizados, enquanto revisões relacionadas a fatores de risco e prognóstico incluem mais frequentemente estudos observacionais comparativos, que são aqueles nos quais o pesquisador coleta dados, simplesmente observando e comparando os eventos enquanto eles acontecem.

Risco refere-se, genericamente, à probabilidade de algum evento indesejado. Nesta diretriz, o termo risco é utilizado num sentido mais restrito, para descrever a probabilidade de que pessoas expostas a certos fatores (fatores de risco) adquiram ou desenvolvam subsequentemente uma doença ou situação de interesse. Os estudos observacionais comparativos de coorte e caso-controle são os delineamentos mais utilizados para avaliação de fatores de risco.

Nos estudos de coorte convencionais, inicialmente um grupo similar de pessoas é dividido em grupos exposto e não exposto a um fator de interesse. Essas pessoas são então observadas prospectivamente por um período de tempo para a avaliação da ocorrência do desfecho pré-determinado. Há ainda os estudos de coorte retrospectivos, que avaliam no presente os desfechos relacionados à uma exposição que ocorreu no passado. Os estudos de coorte retrospectivos geralmente utilizam dados de registros populacionais de determinado país ou região e por isso são mais rápidos que os estudos de coorte prospectivos.

Os estudos de caso-controle são sempre retrospectivos e os indivíduos são agrupados de acordo com a presença (casos) ou não (controles) da doença ou situação de interesse. Por meio da comparação entre estes grupos, procura-se determinar a associação do fator de exposição investigado e a ocorrência do desfecho (doença ou situação) de interesse.

Prognóstico, de maneira simplificada, é a predição do curso futuro de uma doença, após sua instalação. Estudos para avaliação de prognóstico são semelhantes aos estudos de coorte para avaliação de fatores de risco. Inicialmente uma amostra semelhante de pessoas com uma característica em comum (uma doença ou situação) é dividida em grupos exposto e não exposto a um fator de interesse (fator

prognóstico). Essas pessoas são então observadas prospectivamente por um período de tempo para a avaliação da ocorrência do desfecho pré-determinado (por exemplo, progressão da doença já instalada, melhora clínica, outros eventos subsequentes)⁸. Características adicionais sobre os principais tipos de estudos envolvidos em revisões sistemáticas sobre fatores de risco e prognóstico se encontram no Apêndice A.

1.5 Limitações das revisões sistemáticas de estudos observacionais

As principais limitações das revisões sistemáticas e metanálises de estudos observacionais comparativos são decorrentes das limitações do delineamento dos estudos primários (coortes e caso-controles). Basicamente, a presença de variáveis de confusão, ausentes ou controladas em ensaios clínicos randomizados, pode gerar medidas de associação enviesadas nos estudos observacionais. Os estudos primários podem utilizar análises ajustadas, por meio de modelos de regressão multivariados, com o objetivo de diminuir a influência das variáveis de confusão nas medidas de associação. No entanto, esse artifício somente pode ser feito para as variáveis de confusão conhecidas e é limitado ao tamanho do estudo. Assim, as variáveis de confusão não conhecidas ou não controladas nas análises podem gerar estimativas de efeito não fidedignas e as metanálises destes estudos poderão potencializar esse erro.

Por exemplo, estudos observacionais avaliando a terapia de reposição hormonal em mulheres no climatério demonstraram redução no risco de eventos cardiovasculares nessa população. Posteriormente, ensaios clínicos randomizados avaliando a associação da terapia de reposição hormonal e eventos cardiovasculares não confirmaram o benefício previamente demonstrado. A principal explicação para a discrepância entre a evidência gerada pelos estudos observacionais e a gerada pelos ensaios clínicos randomizados foi relacionada com as características inerentes à população exposta ao fator de exposição (terapia de reposição hormonal) nos estudos observacionais. Mulheres climatéricas que optaram por fazer reposição hormonal (ou aquelas cujos médicos indicaram a terapia) apresentaram maior probabilidade de serem expostas também a outros fatores (fatores de confusão) que reduziam o risco de eventos cardiovasculares como, por exemplo, melhor condição socioeconômica, prática de atividade física, realização de exames médicos periódicos, entre outros cuidados gerais com saúde. O benefício da terapia de reposição hormonal não foi demonstrado nos ensaios clínicos randomizados porque as características potencialmente modificadoras do desfecho (fatores de confusão) foram distribuídas aleatoriamente, ficando balanceadas entre os grupos exposto e não exposto. Assim, os grupos gerados no final dos estudos eram verdadeiramente comparáveis, pois apresentaram como única diferença a exposição à terapia de reposição hormonal⁹⁻¹¹.

Groenwold e cols. demonstraram que a maioria dos estudos observacionais publicados em revistas científicas de alto impacto relatam os fatores de confusão controlados nas análises. No entanto, detalhes sobre o processo de seleção e

inclusão destes fatores de confusão são pouco relatados. Além disso, apenas 9% dos estudos observacionais avaliados comentaram sobre o potencial impacto dos fatores de confusão não controlados nas medidas de associação¹².

Adicionalmente, nem todas as informações necessárias à compreensão da condução e resultados dos estudos observacionais são relatadas. Nesse sentido, foi publicada recentemente a diretriz Strengthening the Reporting of Observational studies in Epidemiology (STROBE) com a finalidade de indicar as informações essenciais que devem ser reportadas e como devem ser relatadas, de modo que o leitor da informação possa avaliar criticamente e validar os achados do estudo (Anexo A)^{3, 13}.

1.6 Recursos necessários

Alguns recursos são decisivos para a realização de uma revisão sistemática de forma otimizada e incluem: a) financeiros – frequentemente necessários para obtenção de artigos ou para tradução dos artigos publicados em línguas não dominadas pela equipe; b) tecnologia – como *softwares* estatísticos e gerenciadores de referências e; c) humanos – que devem incluir profissionais com conhecimento clínico sobre o tema, sobre epidemiologia e ter habilidades para manuseio dos recursos tecnológicos. Antes de iniciar uma revisão sistemática, idealmente deve-se responder positivamente às seguintes perguntas:

- 1) Há uma questão clínica bem definida, com elegibilidade bem estabelecida e que ainda não foi respondida?
- 2) Há uma equipe com o mínimo de pessoas treinadas?
- 3) Há tempo suficiente da equipe para se dedicar à condução da revisão de forma a cumprir o cronograma estabelecido?
- 4) Há recursos financeiros disponíveis caso haja necessidade de alguma consultoria específica ou tradução de artigos em língua não dominada pela equipe?
- 5) Há recursos estatísticos (conhecimento, programas específicos) necessários para análise de dados?

A partir destas respostas, deverá ser analisada a possibilidade da aquisição dos recursos não disponíveis. Vale salientar que o tempo é fator decisivo e vai depender muito da quantidade de literatura disponível sobre a questão de pesquisa. O gerenciamento do tempo disponível deve focar nas etapas de planejamento, treinamentos, desenvolvimento de protocolo, busca e seleção dos artigos, elaboração de fichas clínicas, extração e análise dos dados, entre outros. Além da previsão em cronograma do tempo para contato e levantamento de dados com autores cujos artigos não incluem informações necessárias para a revisão sistemática.

Ainda devem ser considerados recursos humanos, financeiros e materiais. Inicialmente será necessário o planejamento do número de pessoas que irão

compor a equipe e suas atribuições. Um coordenador é fundamental para que possa liderar a equipe e reconhecer a necessidade da consultoria de especialistas como bibliotecários, estatístico e/ou profissional para tradução de artigos em idioma que a equipe não domina, além de ser o responsável por consolidar cada uma das etapas de condução da revisão.

Em relação aos recursos financeiros, é altamente recomendável a elaboração de um orçamento detalhado prévio ao início da revisão, que contemple gastos com consultorias, bens materiais permanentes (computadores) e de consumo (material de escritório e impressão, programas estatísticos específicos, se aplicável, programas gerenciadores de referências, entre outros). Esses recursos financeiros podem ser próprios ou financiados por agências de fomento à pesquisa, instituições envolvidas na avaliação de tecnologias em saúde ou na elaboração de diretrizes para prática clínica.

A experiência e conhecimento dos revisores são fundamentais para a garantia da qualidade metodológica. Para tanto, é necessário ter noções básicas de metodologia científica, estatística, epidemiologia, conhecimento clínico sobre a área relacionada à questão de pesquisa, informática e domínio, pelo menos, da língua inglesa. Além disso, a equipe deve saber utilizar as ferramentas necessárias à condução da revisão, como os gerenciadores de referência, programas para cálculo da metanálise, as peculiaridades de busca nas várias bases de dados, além de habilidades de redação de protocolo e do manuscrito final.

2 Planejamento da revisão sistemática

2.1 Definição da questão de pesquisa estruturada no formato do acrônimo PECO

O processo de condução de uma revisão sistemática começa com a apresentação de uma ideia clara e objetiva, independente do delineamento de estudo. É também preciso assegurar que essa questão não tenha sido objeto de pesquisa prévia e com resposta definitiva, o que não justificaria a duplicidade da revisão.

A questão de pesquisa bem formulada envolve extrema especificidade e precisão, tendo em mente a identificação das variáveis e da população de interesse.

O conceito de questão de pesquisa estruturada foi introduzido em 1995 por Richardson e cols.¹⁴ com o intuito de identificar palavras-chave pesquisáveis que representem a questão clínica. A fim de orientar a formulação da questão de pesquisa, convencionou-se estruturá-la segundo os componentes do acrônimo PECO, onde cada letra representa um componente da questão, como apresentado no Quadro 2:

Quadro 2 – Estratégia PECO para formulação da questão de pesquisa

P	População de interesse	Paciente ou problema a ser abordado
E	Exposição	Exposição a ser considerada (por ex.: fator de risco, fator prognóstico)
C	Comparador	Controle (por ex.: não exposição)
O	Desfecho (<i>outcome</i>)	Desfecho de interesse (ou situação de interesse)

Fonte: elaboração própria.

A princípio, os autores sugeriram o uso do acrônimo PICO, onde a letra I representa a intervenção de interesse. A variação sugerida neste documento substitui a letra I por E, onde E refere-se à exposição, substituindo a palavra *intervention* para adaptar a questão à avaliação de risco ou prognóstico. Outros autores propõem o uso do acrônimo PICOS, onde o S refere-se ao desenho de estudo (*study design*). Já, Fineout-Overholt e Johnson¹⁵ sugeriram um esquema de cinco componentes para questões práticas baseadas em evidências, utilizando a sigla PICOT, com T representando o período de seguimento (*timeframe*). É proposta a seguinte abordagem para cada elemento do acrônimo PECO:

P – População: abordar uma população específica e suas características e informações demográficas relevantes, como doença ou condição, estágio/gravidade, idade, gênero, entre outros.

E – Exposição: o fator de exposição deve ser claramente definido, incluindo critérios para sua aferição, dose/quantidade da exposição, entre outros. Esse componente é análogo à definição da intervenção para estudos clínicos randomizados.

C – Controle: Estudos de coorte clássicos (prospectivos) incluem uma comparação com uma população similar que não teve contato com o “fator de exposição”. Em estudos de caso-controle, o grupo controle é um grupo similar, no entanto, sem a presença do desfecho de interesse.

O – Desfecho (*outcome*): definição do desfecho clínico. Se possível, priorizar os desfechos clinicamente relevantes como mortalidade, morbidade, qualidade de vida, etc. No entanto, desfechos substitutos podem fornecer informações relevantes sobre determinada área do conhecimento.

O Quadro 3 exemplifica questões de pesquisa estruturadas para estudos observacionais.

Quadro 3 – Exemplo de questões de pesquisa de estudos observacionais

Qual a incidência de óbitos em 30 dias em pacientes submetidos à cirurgia não-cardíaca e que apresentam troponina T elevada no período perioperatório em relação aos que não apresentam?

P = Pacientes submetidos à cirurgia não-cardíaca

E = Troponina T elevada

C = Sem troponina T elevada

O = Óbito em 30 dias

Fonte: (Devereaux, 2012, baseado)¹⁶.

2.2 Definição dos critérios de elegibilidade

Os componentes da questão de pesquisa no formato PECO auxiliam a definir os critérios de elegibilidade para seleção dos estudos primários. Desenho do estudo, níveis de dose mínima/máxima, o estado de saúde inicial, o período de seguimento mínimo, entre outros, são exemplos de critérios adicionais de elegibilidade.

É importante que essas decisões sejam tomadas em consulta com especialistas da área de interesse. Não recomendamos utilizar nenhum limite na busca, como por exemplo, por data mais recente dos estudos, pois pode resultar na perda de dados originais que, devido aos recursos, à ética ou a outras razões, não tenham sido recentemente publicados.

Assim como para inclusão e exclusão de pacientes nos estudos clínicos primários, os critérios de elegibilidade de estudos em revisões sistemáticas devem ser definidos com antecedência e registrados em protocolo (seção 2.4).

2.3 Revisão de literatura: justificativa para a revisão sistemática

Ao se definir uma questão de pesquisa para a revisão sistemática, podemos nos deparar com a situação em que a associação entre fator de exposição e desfecho já tenha sido estabelecida em estudos prévios.

Assim, deve-se fazer uma breve busca na literatura por uma revisão sistemática já realizada. Para isto, pode-se recorrer ao site oficial da Cochrane Library¹⁷, especializada em revisões sistemáticas e da Biblioteca Cochrane¹⁸, disponível na Biblioteca Virtual em Saúde (BVS), além de fontes adicionais como o Centre for Reviews and Dissemination (CRD)¹⁹.

Outra fonte essencial é o PubMed/MEDLINE por meio do Clinical Queries²⁰ – uma ferramenta da base apresentada na parte média-inferior da página inicial. No Clinical Queries os resultados são categorizados em três grupos: revisões sistemáticas, estudos clínicos ou estudos em genética médica.

Também se recomenda a busca no banco de estudos da Rede Brasileira de Avaliação de Tecnologias em Saúde (Rebrats), o Sisrebrats²¹, que contempla em sua base estudos elaborados pelas instituições membros da Rebrats e estudos financiados através de editais com fomento para pesquisas.

Se for identificada alguma revisão sistemática prévia, abordando a mesma questão de pesquisa de interesse, só se justifica prosseguir na condução do projeto mediante algum diferencial relevante como, por exemplo, a publicação de um novo estudo, que possa modificar as medidas de associação e que não tenha sido considerado na revisão sistemática prévia.

2.4 Documentação da metodologia: redação do protocolo

Assim como em qualquer estudo clínico primário, o protocolo de uma revisão sistemática deve ser elaborado previamente e tem como objetivo registrar todo o processo que envolve a realização da revisão sistemática, bem como definir as análises que serão realizadas. Com isto, o risco de erros sistemáticos ou vieses, introduzidos por decisões que são influenciadas pelos achados, podem ser minimizados.

O protocolo descreve as etapas realizadas na revisão e pode ser estruturado em introdução e justificativa da revisão, objetivos e metodologia. Em metodologia deve-se registrar os critérios de elegibilidade, as bases de dados a serem pesquisadas, a definição da estratégia de busca, o processo de triagem e seleção dos artigos, o processo de extração de dados, o plano de análises estatísticas, os desfechos e as análises de subgrupo e de sensibilidade de interesse. O relato dos resultados, discussão e conclusões fazem parte da redação do manuscrito. Sugere-se que o protocolo da revisão seja elaborado, preferencialmente, por um grupo de revisores com experiência tanto na área do conhecimento clínico, quanto em metodologia de revisões sistemáticas. O conhecimento da área de investigação é fundamental para a definição prévia das variáveis de confusão e fatores de interação que serão considerados nas análises, sob o risco de encontrar falsas associações significativas, decorrentes do aumento do erro tipo *alfa*.

Uma sugestão de modelo de protocolo está disponível no programa Review Manager, que estrutura o protocolo de revisão sistemática nos moldes exigidos para publicação da revisão na Cochrane Library. O Review Manager está disponível para *download* gratuito no seguinte *website*: <<http://tech.cochrane.org/revman/download>>. O Quadro 4 apresenta a estrutura de protocolo sugerida pela Cochrane Library²².

Quadro 4 – Estrutura de protocolo sugerida pela Cochrane Library

Título em linguagem simples

Introdução

Sugere-se estruturar da seguinte forma:

- Descrição da condição de saúde de interesse
- Descrição da exposição
- Como a exposição pode afetar a condição de saúde de interesse
- Por que é importante conduzir esta revisão

Objetivos

Métodos

Sugere-se estruturar da seguinte forma:

- Critérios para considerar estudos para esta revisão (Elegibilidade)
 - Tipos de estudos
 - Tipos de participantes
 - Tipos de exposição
 - Tipos de desfechos
- Estratégias de busca para identificação dos estudos
 - Busca em bases eletrônicas
 - Busca em outras fontes
- Coleta e análise de dados
 - Seleção de estudos
 - Extração e gerenciamento de dados
 - Avaliação do risco de viés nos estudos incluídos
 - Avaliação da heterogeneidade
 - Avaliação do viés de publicação
 - Síntese de dados
 - Análises de subgrupo e investigação de heterogeneidade
 - Análises de sensibilidade

Referências

Fonte: (Higgins, 2011, baseado)²².

Desde 2005, editores dos mais prestigiados periódicos médicos passaram a exigir o registro do protocolo dos ensaios clínicos em bases de dados públicas, antes da condução do estudo²³. Esta medida visa garantir que o estudo seja conduzido de acordo com o plano inicial do protocolo, independente dos resultados obtidos serem positivos ou não. Para a condução de revisões sistemáticas e metanálises, entretanto,

esse registro ainda não é exigido pela maioria dos periódicos. Apesar de não obrigatório, protocolos de revisões sistemáticas e metanálises podem ser registrados no International Prospective Register of Ongoing Systematic Reviews (PROSPERO)²⁴.

PROSPERO é uma base de dados internacional que permite o registro prospectivo de revisões sistemáticas na área da saúde e assistência social. Tem como objetivos adicionais fornecer uma lista detalhada de revisões sistemáticas registradas para ajudar a evitar a duplicação não planejada, além de permitir identificar se os resultados da revisão publicados estavam de acordo com o plano original de trabalho.

DIRETRIZ: PLANEJAMENTO DA REVISÃO SISTEMÁTICA

- 1) Formular questão de pesquisa estruturada de acordo com o acrônimo PECO**
- 2) Buscar revisões prévias na literatura**
- 3) Redigir protocolo**

3.1 Busca de potenciais estudos elegíveis

A busca por estudos observacionais deve ser abrangente o suficiente para assegurar a recuperação de todas as evidências disponíveis para a questão de pesquisa e, acima de tudo, deve ser reproduzível, para garantir sua validade. A ajuda de um bibliotecário especialista em elaborar estratégias de busca ou um Trials Search Coordinator (TSC) para revisões sistemáticas geralmente é de grande auxílio na condução da revisão.

Estratégias de busca para identificar estudos observacionais são mais bem elaboradas se guiadas pelos domínios do acrônimo PECO. Conforme mencionado anteriormente, os delineamentos de estudos observacionais mais utilizados para avaliação de risco e prognóstico são os estudos de coorte e caso-controle. Estabelecer os termos de busca para estes delineamentos não é tarefa fácil, por conta de inconsistências e diversidades da terminologia usada tanto no texto da publicação quanto na indexação nas várias bases de dados.

A busca por evidências deve contemplar as bases de dados consideradas essenciais, como MEDLINE, EMBASE, e Latin American and Caribbean Health Sciences Literature (LILACS)²⁵. Nesse sentido, Golder e Whiting (2006) demonstram que a busca somente pelo MEDLINE não é considerada adequada para revisões sistemáticas, levando à perda de estudos relevantes, não indexados na base²⁶. Também devem ser investigadas as fontes de dados não publicados e literatura cinzenta, que serão mais bem explorados na sequência. Além disso, pode-se incluir a busca manual pela leitura de anais e resumos de congressos na área de interesse, listas de referências de estudos publicados, demais referências relevantes, por contato com especialistas, etc. Quanto mais abrangente for a busca por estudos, menor o risco de viés de publicação na revisão sistemática. A descrição mais detalhada das fontes de evidências segue nas próximas seções.

3.1.1 Bases de dados bibliográficas eletrônicas

A busca por potenciais estudos sempre se inicia pelas grandes bases de dados eletrônicas como MEDLINE e EMBASE. Também se recomenda pesquisar na LILACS, SCOPUS, Web of Science²⁷, Cumulative Index to Nursing and Allied Health Literature (CINHAL)²⁸, Australasian Medical Index, Chinese Biomedical Literature Database, outras bases de dados especializadas nacionais e regionais, anais de congressos, bancos de teses, etc.

Nem todas as bases de dados eletrônicas são de livre acesso. Algumas bases permitem livre acesso, porém, limitado à citação e ao resumo da publicação.

O MEDLINE é uma das mais importantes bases de dados internacional. Contém mais de 23 milhões de citações de resumos e referências de artigos no campo da medicina, biomedicina, ciências da vida e áreas correlatas. O MEDLINE pode ser acessado via portal do PubMed, um serviço da Biblioteca Nacional de Medicina dos Estados Unidos. A base contém citações datadas a partir de 1948 e disponibiliza, em alguns casos, *links* para acesso ao texto completo nos *websites* das editoras dos periódicos. Pode ser acessado também por outras plataformas, como o OVID (acesso via assinatura) e a BVS, de livre acesso.

O EMBASE é focado na publicação da literatura européia, atualmente com mais de 27 milhões de citações. É fundamental que a busca contemple também o EMBASE. No artigo de Wilkins e cols.²⁹ intitulado “Embase *versus* Medline searches, in Family medicine” fica claro a importância da busca no EMBASE. Os autores comparam os resultados de busca no MEDLINE e EMBASE em 15 tópicos selecionados e verificam que somente 5% das referências recuperadas foram listadas nas duas bases. Também observam que a busca no EMBASE recuperou mais artigos que no MEDLINE.

A LILACS, coordenada pela BVS, é o mais importante e abrangente índice da literatura científica e técnica da América Latina e do Caribe, composta por 31 países. Uma vantagem é que uma grande parte dos artigos está disponível em forma de texto completo por meio da Scientific Electronic Library Online (SciELO)³⁰.

Teses e dissertações normalmente não estão indexadas nas grandes bases de dados bibliográficas, como MEDLINE ou EMBASE, mas devem ser igualmente pesquisadas. Nesse sentido, as bases de dados de teses e dissertações também são de grande auxílio. No Brasil, dispomos de alguns bancos de teses coordenados por grandes universidades como a Universidade de São Paulo (USP), a Pontifícia Universidade Católica (PUC), Portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)³¹, disponibilizado pelo Ministério da Educação e no Instituto Brasileiro de Ciência e Tecnologia (IBICT). Ainda, é possível pesquisar em bases internacionais de teses e dissertações, como por exemplo, a ProQuest Dissertations Theses Database³².

O Google Acadêmico é uma boa ferramenta de busca, pois permite integrar várias fontes de pesquisa em um só lugar e tem a função de complementar a busca formal. A busca pode recuperar artigos revisados por especialistas, editores, teses, livros, resumos e artigos de editoras acadêmicas, organizações profissionais, bibliotecas de pré-publicações, universidades e outras entidades acadêmicas. A Turning Research into Practice database (TRIP)³³ também é uma ferramenta de busca e traz os resultados por categorias como: diretrizes, revisões, artigos do MEDLINE, informações para pacientes, capítulos de livros.

3.1.2 Pesquisa em fontes de dados não publicados, em andamento e literatura cinzenta

Muitos estudos são finalizados, mas não são publicados. As principais causas dessa associação seriam resultados não significativos ou negativos. A identificação destes estudos não é fácil. Uma possível maneira de rastreá-los é por meio da busca em bases de dados que registram protocolos de estudos. São exemplos de bases de registros de protocolo: ClinicalTrials.gov³⁴, International Standard Randomised Controlled Trial Number Register (ISRCTN)³⁵, nos EUA, o EU Clinical Trial Register, na Europa e o Registro Brasileiro de Ensaios Clínicos (ReBEC)³⁶. Apesar destas bases terem sido criadas para o registro de ensaios clínicos, estudos observacionais comparativos em andamento ou finalizados também podem ser encontrados. O limitante dessa estratégia é que estudos menores, menos relevantes, podem não estar registrados nessas bases.

Se identificado um protocolo de estudo registrado com *status* de finalizado e não for identificada a publicação correspondente, esforços devem ser realizados para contatar os autores, a fim de obter os dados do estudo.

A Figura 1 exibe um exemplo de registro de protocolo no ClinicalTrials.gov.

Figura 1 – Exemplo de registro de protocolo no ClinicalTrials.gov
“The brain natriuretic peptide as a predictor of mortality in cardiac surgery”.

The screenshot displays a ClinicalTrials.gov study record. At the top, it indicates 'Trial record 42 of 351 for: "brain natriuretic peptide"' with navigation links for 'Previous Study', 'Return to List', and 'Next Study'. The study title is 'Evaluation of the Brain Natriuretic Peptide as a Predictor of Morbidity and Mortality in Cardiac Surgery'. A status message states: 'The recruitment status of this study is unknown because the information has not been verified recently.' It is verified by 'University Hospital, Strasbourg, France' as of May 2010, with a recruitment status of 'Recruiting'. The sponsor is 'University Hospital, Strasbourg, France'. The clinicalTrials.gov identifier is 'NCT00699673'. The record shows it was first received on June 10, 2008, last updated on May 21, 2010, and last verified on May 2010. A link for 'History of Changes' is provided. Navigation tabs include 'Full Text View', 'Tabular View', 'No Study Results Posted', 'Disclaimer', and 'How to Read a Study Record'. The 'Purpose' section describes the objective: to assess if perioperative variation of Brain Natriuretic Peptide (BNP) levels is a predictor of mortality and morbidity after cardiac surgery. It details the material and methods, including 500 consecutive patients enrolled prospectively before cardiac surgery under cardiopulmonary bypass, with BNP levels measured pre-surgery and on postoperative day 1. The hypothesis states that perioperative BNP variations may be more sensitive than pre- or postoperative BNP levels alone, and that perioperative homeostasis will be measured to assess its impact on BNP secretion. A table lists the condition as 'Coronary Artery Disease' and 'Valvular Disease'. The study type is 'Observational', the design is 'Observational Model: Case-Only', and the time perspective is 'Prospective'. The official title is 'Evaluation of the Brain Natriuretic Peptide as a Predictor of Morbidity and Mortality in Cardiac Surgery'.

Fonte: ClinicalTrials.gov.

A literatura cinzenta é aquela não indexada em periódicos que publicam formalmente a literatura científica. Segundo o *Handbook* da Cochrane, 2011 (cap. 6, seção 6.2.1.8), este tipo de literatura também deve ser considerado no processo de busca das evidências, pois, consiste na fonte de aproximadamente 10% de estudos referenciados pelas revisões da Cochrane. Nesse sentido, anais de congressos são importantes fontes de literatura cinzenta, disponibilizados por algumas revistas de especialidades em edições especiais ou suplementos. Assim, este tipo de literatura pode estar disponível na mesma base de dados eletrônica onde a edição regular da revista está indexada. Como exemplos também temos a Web of Science que indexa *abstracts* de Congressos, o BMC Central Proceedings e o BMC Meeting Abstracts^{37, 38}.

O acesso à literatura cinzenta pode também ser feito por busca manual, página-a-página de cada artigo de revistas da área de interesse que não estão indexadas nas grandes bases de dados eletrônicas. Existem ainda bases de dados específicas que cobrem a literatura cinzenta como o OpenGrey (anteriormente chamado de OpenSIGLE) sistema de informação da literatura cinzenta Européia e a National Technical Information Service (NTIS) que oferece acesso aos resultados de pesquisas governamentais e não governamentais dos Estados Unidos e estão disponíveis via internet.

Além dessas fontes alternativas de estudos, contato com colegas pesquisadores, especialistas no assunto da revisão, autores que se destacam em estudos que foram selecionados ou empresas de interesse na área a ser pesquisada devem ser consultados, como fonte adicional de informação.

Por último, é recomendado verificar a lista de referências dos estudos incluídos na revisão, a fim de detectar alguma evidência relevante que possa não ter sido recuperada pela estratégia de busca.

3.1.3 Elaboração da estratégia de busca nas várias bases de dados

A elaboração da estratégia de busca é a componente chave da revisão. A definição dos termos apropriados de busca deve priorizar a sensibilidade em detrimento da especificidade, garantindo o resgate de toda a potencial evidência disponível. O conhecimento dos mecanismos de busca nas várias bases de dados se faz necessário, já que particularidades das bases as tornam diferentes entre si.

Alguns detalhes, se bem observados, ajudarão na construção da estratégia. Geralmente, a construção da estratégia de busca em revisões sistemáticas de estudos observacionais é baseada nos domínios do acrônimo PECO da questão de pesquisa. Para ensaios clínicos randomizados, recomenda-se construir a estrutura considerando termos para o P (população), I (exposição/intervenção), e C (controle), não limitando a busca para os desfechos, sob o risco de não recuperar estudos elegíveis. No caso de estudos observacionais, os termos que caracterizem o desfecho são essenciais, principalmente quando o desenho de estudo é caso-controle, em

que a seleção dos casos parte da presença do desfecho. Por outro lado, termos que caracterizem o controle não são importantes, pois muitas vezes, o controle é a situação de não exposição ao fator em estudo.

3.1.3.1 Seleção dos termos para a busca

Sempre que possível, deve-se utilizar o vocabulário controlado, que é o descritor de assunto no qual o artigo foi indexado na base. No MEDLINE, este vocabulário é denominado Medical Subject Heading, MeSH, no EMBASE, Emtree; e na LILACS, os Descritores em Ciências da Saúde (DeCS).

Sugerimos que a seleção dos termos sempre se inicie pelos descritores do MeSH no MEDLINE. Para a busca no EMBASE, deve-se identificar o Emtree correspondente ao MeSH, já que nem sempre um termo MeSH corresponde ao mesmo termo no Emtree. Entretanto, para a maioria dos termos MeSH, o EMBASE já sugere o termo correspondente ao Emtree. Abaixo, um exemplo de como o mesmo termo descritor de assunto pode variar nas bases:

Ex.: Termo no MeSH: *Natriuretic Peptide, Brain*

Ex.: Termo no Emtree: *natriuretic peptide, brain use: brain natriuretic peptide*

Na base LILACS, a definição do vocabulário controlado é feita pelo DeCS que é a tradução do MeSH. Para aquele revisor que tenha dúvida quanto ao termo em inglês, o DeCS é uma boa opção, pois oferece a busca nas três línguas da base: inglês, português e espanhol. Recomenda-se que a estratégia de busca seja elaborada com a “soma” dos termos nas três línguas por meio do operador booleano “OR”. O Apêndice C ilustra um exemplo desta estratégia.

Um recurso também muito útil na hora da seleção dos termos, principalmente no MEDLINE, é quando os autores já possuem um artigo sobre o assunto e que o mesmo está indexado no MEDLINE. Assim, é possível identificar quais os termos MeSH a base utilizou para indexar aquele artigo e com isto construir a estratégia de busca mais adequada para a questão a ser estudada na revisão.

A estratégia de busca não deve ficar restrita somente aos descritores de assunto. Sugere-se que a estratégia seja incrementada com o uso de vocabulário não controlado que consiste em palavras de texto, sinônimos, siglas, termos relacionados, palavras chave e variações de grafia.

Na base MEDLINE, o vocabulário não controlado corresponde aos “*entry terms*” que se encontram dentro da definição do termo MeSH. Estes termos representam sinônimos, indexações prévias ou derivações do assunto, que contribuem para sensibilização da estratégia. A quantidade destes termos varia de acordo com o termo MeSH. No Apêndice B encontra-se um exemplo de estratégia sensibilizada da busca.

Para a elaboração da estratégia de busca completa, sugere-se a definição dos termos de busca para cada domínio relevante da questão de pesquisa. De forma geral, os domínios relevantes para a busca por estudos observacionais são: P (população), E (exposição) e O (desfecho).

Os termos relacionados a um mesmo domínio (por ex.: todos os termos referentes à população) devem ser “somados” por meio do operador booleano “OR”. Após, os domínios (por ex.: termos referentes à população e termos referentes à exposição) devem ser “relacionados” por meio do operador booleano “AND”. Uma explicação mais detalhada da construção da estratégia de busca e utilização dos operadores booleanos encontra-se no Apêndice C.

3.1.3.2 Filtros de busca para estudos observacionais

Filtros de busca são estratégias de busca validadas para recuperar determinado tipo de estudo. O uso de filtros para busca por estudos observacionais ainda é uma recomendação controversa. Existem inconsistências no vocabulário de indexação destes estudos nas diferentes bases de dados e muitas vezes o descritor indexado não corresponde ao delineamento de fato empregado. Assim, o uso de filtros de busca pode acarretar na perda de estudos potencialmente relevantes para a revisão sistemática.

Existem iniciativas de grandes centros para o desenvolvimento de filtros para busca de estudos com foco no desenho empregado ou área do conhecimento, como a McMaster University, através do Health Information Research Unit (HIRU) e a do The InterTASC Information Specialists’ Sub-Group Search Filter Resource do Centre for Reviews and Dissemination (CRD).

O PubMed Clinical Queries³⁹ apresenta os filtros elaborados por Haynes e cols.⁴⁰ para vários desenhos de estudos, entre eles estudos de etiologia e prognóstico. Caso os revisores optem pela utilização de algum filtro, o ideal é que utilizem filtros já publicados e validados. No entanto, ressaltamos que ainda não existe consenso sobre o uso ou não de filtro para recuperação de estudos observacionais para revisão sistemática.

3.1.4 Relatando o processo de busca

Deve-se descrever em protocolo e relatar na seção de metodologia do manuscrito a lista de todas as bases de dados que serão pesquisadas, com suas respectivas datas de abrangência (por exemplo: MEDLINE de 1950 a 31 de janeiro de 2014) e datas de acesso. Demais fontes de busca, como literatura cinzenta, contatos, busca manual, etc., também devem ser relatadas.

As estratégias de busca utilizadas para cada base de dados devem ser disponibilizadas no protocolo ou como apêndices do manuscrito da revisão. É importante este registro para futuras atualizações da revisão sistemática e para garantir a reprodutibilidade da busca.

DIRETRIZ: ESTRATÉGIA DE BUSCA

1. Buscar evidência disponível nas principais bases de dados bibliográficas eletrônicas: MEDLINE, EMBASE e LILACS.
2. Buscar evidências provenientes de literatura cinzenta.
3. Definir quais termos serão utilizados para os itens P, E e O da questão de pesquisa no formato PECO, lembrando que, para estudos observacionais, o desfecho é relevante para a busca.
4. Elaborar a estratégia de busca utilizando vocabulário controlado e vocabulário livre, juntamente com os operadores booleanos para as combinações entre os termos.
5. Utilizar filtros para estudos observacionais, em casos específicos, particularmente se a busca recuperar quantidade de citações inviável para a avaliação. Neste caso, utilizar filtros publicados pelo HIRU, CRD ou PubMed.

3.2 Avaliação da elegibilidade dos estudos

O processo de avaliação da elegibilidade passa por uma etapa de triagem dos artigos, com leitura de título e resumo (quando disponível) e uma etapa de confirmação, pela leitura do texto completo. As etapas estão descritas a seguir:

3.2.1 Triagem dos artigos pela leitura de título e resumo

A busca de estudos em todas as possíveis fontes de dados gera um número muito maior de artigos do que os que realmente serão elegíveis pelos critérios estabelecidos. Isto ocorre porque a estratégia de busca é elaborada preconizando a sensibilidade em detrimento à especificidade, ou seja, a estratégia recupera muito mais artigos caracterizados como “falso positivos” (não elegíveis para a revisão sistemática), porém reduz ao mínimo a probabilidade de não recuperar um artigo elegível.

Quando se trata de triagem de ensaios clínicos randomizados para inclusão em uma revisão sistemática, a leitura rápida do título e resumo permite descartar um grande número de referências que não se enquadram nos critérios de elegibilidade estabelecidos pela revisão. No entanto, quando o escopo são estudos observacionais comparativos, nem sempre é possível descartar o estudo após leitura do resumo. Isso se deve à rotulação inadequada do desenho de estudo e, por isso, frequentemente é necessária a leitura do texto integral para verificação do desenho de estudo de fato empregado.

O grupo de métodos de estudos não randomizados da Colaboração Cochrane⁴¹ recomenda que o tipo de estudo seja definido pelas características do desenho empregado, em vez do rótulo do tipo de estudo descrito na referência. Nesse sentido,

esse grupo propôs um “*checklist*” com itens que podem auxiliar na identificação do desenho de estudo. Estes itens buscam identificar características-chave que, com base na experiência do grupo, definem a maioria dos estudos observacionais. Assim, sugerimos nessa diretriz uma adaptação deste “*checklist*”, que pode ser utilizado pelos autores da revisão nos casos de dúvidas quanto ao desenho do estudo primário (Tabela 3). No entanto, reforçamos que o conhecimento prévio sobre epidemiologia e tipos de estudos observacionais é fundamental para a correta identificação do estudo.

Quadro 5 – Características presentes em estudos observacionais

	Coorte prospectiva	Coorte retrospectiva	Caso-controle
Os participantes foram alocados nos grupos por:			
Randomização sigilosa?	Não	Não	Não
“Quase” randomização (quasi-randomization)?	Não	Não	Não
Por outra ação dos pesquisadores?	Não	Não	Não
Preferência dos participantes?	Possivelmente	Possivelmente	Possivelmente
Com base no desfecho?	Não	Não	Sim
Qual parte do estudo foi prospectiva:			
Identificação dos pacientes?	Sim	Não	Não
Verificação das características de base e alocação da exposição	Sim	Não	Não
Avaliação dos desfechos	Sim	Possivelmente	Não
Os grupos foram comparáveis para as variáveis:			
Potenciais variáveis de confusão?	Possivelmente	Possivelmente	Possivelmente
Avaliação de base das variáveis relacionadas ao desfecho?	Possivelmente	Possivelmente	Não

Fonte: (Higgins e cols., 2011, adaptado)⁴¹.

O uso de um *software* de gerenciamento de referências para a triagem dos artigos não é obrigatório, no entanto, as facilidades desta ferramenta em relação à organização das referências, praticidade e otimização de tempo são inquestionáveis. Existem inúmeros programas gratuitos e comerciais, nas mais diversas plataformas de acesso. O Apêndice D fornece mais informações sobre essas ferramentas.

O resultado da soma das citações recuperadas em todas as bases de dados deve ser registrado a fim de elaborar o fluxograma de seleção dos artigos. Ainda, um mesmo artigo pode estar indexado em mais de uma base de dados, de modo que a busca pode recuperar referências duplicadas. O número de citações duplicadas também deve ser apresentado como resultados no fluxograma de seleção dos artigos (Anexo B). Após o registro do número de citações encontradas pelas estratégias de busca e exclusão das citações duplicadas, as remanescentes devem ser distribuídas entre os revisores.

É necessário, no mínimo, uma dupla de revisores para realizar a triagem pela leitura de título e resumo, de maneira independente. Ou seja, revisor A e B devem realizar a triagem das mesmas referências. Se o número de citações recuperadas pela estratégia de busca for muito amplo, ou o tempo disponível para a realização da tarefa for escasso, frequentemente mais de uma dupla de revisores realizam a tarefa. Ou seja, revisores A e B realizam a triagem da primeira metade das referências, enquanto revisores C e D triam a segunda metade das referências.

Na fase de triagem, os revisores devem excluir os artigos que claramente não preencham os critérios de elegibilidade e manter os que são possivelmente elegíveis. Assim, nos casos em que houver dúvidas sobre a relevância da citação, por parte de pelo menos um dos revisores, a referência deve permanecer entre as potencialmente elegíveis e passar para a etapa seguinte, de avaliação da elegibilidade pela leitura do texto completo. Quando o resumo do artigo não estiver disponível, mas o título for sugestivo de inclusão, também se preconiza manter a referência entre as potencialmente elegíveis.

As discordâncias entre os revisores devem ser resolvidas preferencialmente por um terceiro, mas em algumas situações os dois revisores podem chegar a um consenso sem a necessidade do terceiro. Em alguns casos, pequenas divergências entre os revisores podem ser desconsideradas, visto que estas referências terão sua elegibilidade confirmada na etapa seguinte. Por outro lado, se o resultado entre a dupla for muito divergente, vale a pena que os revisores resolvam as discordâncias, de forma a otimizar o tempo na recuperação dos textos completos de artigos. A concordância entre os revisores pode ser mensurada usando a estatística Kappa de Cohen. Quanto mais próximo de um o valor de Kappa encontrado, maior é a concordância entre os revisores. Um valor de Kappa igual ou superior a 0,6 é aceitável. No entanto, a avaliação formal da concordância por meio de teste estatístico não é mandatória.

É importante ressaltar que a unidade de análise de uma revisão sistemática é o estudo e não o artigo ou a citação/referência. Um único estudo pode gerar mais de uma citação/referência (que pode ser um artigo, um pôster apresentado em congresso, um resumo publicado nos anais de congresso, etc.). De modo semelhante, um único artigo pode relatar mais de um estudo. Atenção deve ser dada quanto a estas diferenças para não causar confusão no momento de incluir um estudo na revisão.

3.2.2 Confirmação da elegibilidade: leitura do artigo completo

Todos os estudos que foram selecionados na fase anterior devem ter sua elegibilidade confirmada pela leitura integral do artigo. Assim como na etapa de triagem, a confirmação da elegibilidade é realizada por duplas de revisores, de modo independente.

A etapa de confirmação da elegibilidade é, geralmente, onde os revisores empregam a maior parte do tempo previsto para a realização da revisão. Conforme citado anteriormente, a categorização do desenho do estudo deve ser realizada com base nas

características do desenho, em vez do “rótulo” do estudo, que muitas vezes só pode ser definitiva após leitura do texto completo.

Nessa etapa, a razão primária da exclusão deve ser registrada para compor o fluxo de seleção dos artigos (Anexo B). Para guiar essa etapa, utiliza-se uma ficha de avaliação da elegibilidade padronizada elaborada previamente, que contém basicamente os critérios de elegibilidade (inclusão e exclusão) estabelecidos. Esta ficha deve conter uma folha de rosto com a identificação do título da revisão (um mesmo revisor pode realizar mais de uma revisão sistemática simultaneamente), nome do revisor, identificação, ou acrônimo do estudo, autor e jornal. O Apêndice E apresenta um modelo de formulário de avaliação da elegibilidade.

Havendo discordância entre os revisores, estas devem ser resolvidas ou por consenso ou por um terceiro. Após, obtêm-se o total de estudos que são, de fato, elegíveis e que irão compor a revisão sistemática. A concordância entre os revisores pode ser mensurada usando a estatística Kappa de Cohen⁴².

DIRETRIZ: AVALIAÇÃO DA ELEGIBILIDADE DOS ARTIGOS

- 1) Somar os resultados de busca de todas as bases.
- 2) Preferencialmente, utilizar um gerenciador de referências para avaliação da elegibilidade.
- 3) Remover artigos duplicados.
- 4) Triar os artigos pela leitura de título e resumo (quando disponível).
- 5) A triagem dos artigos deve ser realizada por dupla de revisores, de forma independente. As discordâncias podem ser resolvidas por consenso ou por meio de um terceiro revisor.
- 6) Confirmar a elegibilidade dos artigos triados pela leitura do texto completo do artigo.
- 7) Na etapa de confirmação da elegibilidade, utilizar ficha de avaliação padronizada contendo os critérios de elegibilidade, a fim de registrar os motivos de exclusão nesta etapa.
- 8) A confirmação da elegibilidade deve ser realizada por dupla de revisores, de forma independente.
- 9) Resolver as discordâncias por consenso ou por meio de um terceiro revisor.
- 10) Se desejado, aplicar o teste estatístico Kappa para quantificar a concordância entre os revisores.

3.3 Extração de dados

3.3.1 Processo de extração de dados

A extração dos dados dos estudos primários incluídos na revisão sistemática é sempre guiada por uma ficha padronizada, que pode ser construída em planilha eletrônica ou como os revisores julgarem mais adequado e de fácil manuseio. Assim como na etapa de seleção dos artigos, a extração de dados também deve ser realizada por dupla de revisores, de maneira independente.

Uma vez extraídos os dados, as informações são confrontadas entre os revisores. Havendo discordâncias nos dados coletados, estas podem ser resolvidas por consenso entre a dupla ou por consulta a um terceiro revisor.

Para extração correta e padronizada dos dados, sugere-se um treinamento prévio entre os revisores que participarão da coleta de dados, de forma a uniformizar os conceitos sobre o tema abordado na revisão.

Nesta fase, relembramos que a unidade de análise de uma revisão sistemática é o estudo e não o artigo ou a publicação/referência. Assim, deve ser feito um formulário de extração de dados para cada estudo incluído e não para cada publicação/referência/artigo. Se um único estudo gerou mais de uma publicação, o formulário deste estudo poderá conter os dados relatados em mais de uma publicação.

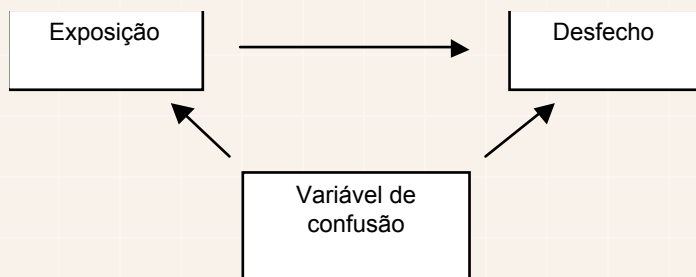
3.3.1.1 Elaboração da ficha de extração de dados

A ficha de extração de dados deve permitir coletar dados de todas as variáveis que forem consideradas importantes para interpretação, cálculo e aplicabilidade dos resultados. Para revisões sistemáticas de estudos observacionais em particular, são informações de interesse:

- tipo de estudo (definido pelas características do desenho empregado);
- critérios de definição da exposição de interesse;
- fatores de confusão considerados;
- características que possam ser potenciais fatores de interação;
- comparabilidade dos grupos em relação aos fatores de confusão considerados;
- métodos considerados para controle dos fatores de confusão;
- critérios de definição do desfecho de interesse;
- medidas de associação da análise bivariada e multivariada (se disponíveis).

O principal limitante na interpretação dos resultados de estudos observacionais é a presença dos fatores de confusão. Devem ser consideradas como fatores de confusão todas as variáveis que estiverem associadas com a exposição e com o desfecho, mas que não estejam na via da cadeia causal entre eles (Figura 2). Por exemplo, um estudo realizado para avaliar o risco de infarto do miocárdio em indivíduos que consomem cinco ou mais doses de café por dia deveria considerar como potencial variável de confusão o fato do paciente ser ou não tabagista. Ser tabagista é considerado um fator de confusão, pois esta variável pode estar associada tanto ao consumo de café, quanto ao risco de infarto do miocárdio. Assim, na análise estatística seria possível “isolar” o efeito do tabagismo na ocorrência do infarto, restando somente a medida de associação da exposição e desfecho (café e infarto).

Figura 2 – Apresentação esquemática da variável de confusão e sua relação com exposição e com desfecho



Fonte: elaboração própria.

Nem sempre é simples a identificação das potenciais variáveis associadas tanto à exposição quanto ao desfecho. É importante fazer uma revisão da literatura abrangente incluindo a leitura de revisões narrativas ou capítulos de livros clássicos para identificar todos os potenciais fatores de risco e outras variáveis que podem contribuir para a ocorrência do desfecho em questão. Essa instrução prévia também auxiliará nas etapas seguintes da revisão sistemática, onde os revisores precisarão avaliar criticamente e verificar se o estudo que está sendo avaliado incluiu todos os potenciais fatores de confusão em relação ao desfecho. Complementando, muitas vezes é necessário consultar epidemiologistas ou especialistas na área relacionada à exposição e/ou ao desfecho de interesse, a fim de contemplar todas as variáveis que possam ser fatores de confusão.

Variáveis de interação também influenciam na interpretação dos resultados gerados por estudos observacionais. Variáveis de interação são aquelas que modificam a medida de associação entre a exposição e o desfecho, entretanto, sua ausência não elimina a associação existente, apenas modifica a magnitude da associação. Por exemplo, a combinação de tabagismo e consumo de álcool é fator de risco para câncer de esôfago (*odds ratio* 8,9; intervalo de 95% de confiança 6,2–13,0). Porém, quando se avalia o efeito do tabagismo em indivíduos que não consomem álcool, a associação com a doença ainda permanece (*odds ratio* 3,0; intervalo de 95% de confiança 1,5–6,2). Assim, o consumo de álcool tem papel de variável de interação, modificando a medida de associação⁴³. O Quadro 6 resume as principais variáveis que devem compor a ficha de extração de dados.

Quadro 6 – Variáveis consideradas para elaboração da ficha de extração dos dados

Desenho do estudo

Características da amostra selecionada:

- Local do estudo e período de recrutamento de pacientes;
- Critérios de inclusão e exclusão;
- Em estudos de coorte:
 - período de seguimento;
 - número de participantes no grupo exposto e não exposto;
 - número de participantes analisados em cada grupo;
 - contaminação (pacientes não expostos sendo expostos).
- Em estudos de caso-controle:
 - critérios para o pareamento dos controles;
 - características.

Variáveis independentes:

- Fatores de exposição sob investigação com seus respectivos critérios de definição ou diagnóstico;
- Potenciais fatores de confusão. Em geral, são as variáveis consideradas nas análises multivariadas, no entanto, podem ser coletadas outras variáveis que, segundo julgamento do revisor, possam ser variáveis de confusão, mesmo que os autores não as tenham considerado nos modelos de análise multivariada.

Variáveis dependentes: desfechos avaliados

- Desfechos avaliados com seus respectivos critérios de definição ou diagnóstico;
- Para escalas: limite superior e inferior; categorizações (se aplicável);
- Em estudos de coorte:
 - os desfechos foram medidos da mesma forma nos grupos comparados?
 - número de participantes expostos com desfecho e número de participantes não expostos com desfecho;
- Em estudos de caso-controle:
 - número de casos e número de controles.

Métodos

- Métodos estatísticos utilizados (modelos de análises, por exemplo: escore de propensão ou regressão logística e variáveis ajustadas);
- Participantes incluídos consecutivamente ou não;
- Esquema de cegamento (se aplicável):
 - de avaliadores de desfecho.

Fonte: elaboração própria.

3.3.1.2 Características dos estudos incluídos e síntese qualitativa dos dados

É necessário apresentar uma tabela de “Características dos estudos incluídos” na seção de resultados de revisões sistemáticas. Esta tabela permite a comparabilidade das variáveis entre os estudos que podem afetar a magnitude da medida de associação entre exposição e desfecho. Também auxilia na interpretação crítica e validação da aplicabilidade dos resultados. No caso de estudos observacionais, essa importância é ainda maior, considerando que muitas dessas diferenças clínicas podem ser

consideradas fatores de confusão e será necessário identificar como e se o estudo primário considerou a presença desses fatores.

A variação das características de base dos pacientes entre os estudos é chamada de heterogeneidade clínica e pode ser fonte de inconsistência entre os achados, principalmente em metanálises de estudos observacionais.

Em metanálises de ensaios clínicos randomizados, se a análise da tabela que contém essas características (geralmente numerada como “Tabela 1”) sugere diferenças clínicas nos pacientes incluídos nos diferentes estudos, mas as análises estatísticas que permitem quantificar esta heterogeneidade não demonstram significância, podemos, então, explorar os achados, no sentido de atribuir maior validade externa à intervenção. Essa assertiva decorre do fato que, em ensaios clínicos randomizados, tanto os fatores de confusão conhecidos quanto os fatores de confusão desconhecidos são uniformemente distribuídos entre os grupos, ou seja, os grupos são verdadeiramente comparáveis entre si e a única diferença entre os grupos pode ser atribuída à intervenção. No caso dos estudos observacionais, algumas variáveis serão consideradas como fatores de confusão e não necessariamente todos os fatores conhecidos serão contemplados nem uniformemente considerados entre os diferentes estudos. Adicionalmente, existem outros inúmeros fatores de confusão que não são conhecidos no momento em que o estudo foi conduzido, como por exemplo, variabilidade genética. A randomização permite eliminar esse tipo de viés. Em metanálise de estudos observacionais, os fatores de confusão e de interação considerados podem ter impacto direto na medida de associação.

3.3.1.3 Avaliação do risco de viés dos estudos incluídos

A avaliação do risco de viés dos estudos incluídos em uma revisão sistemática é importante, independentemente do desenho do estudo primário. Particularmente em estudos observacionais comparativos, a interpretação dos resultados da revisão sistemática pode ser mais problemática devido ao maior número de fontes de vieses.

Da mesma forma que para os ensaios clínicos randomizados, diversas ferramentas para avaliação do risco de viés estão disponíveis e não há um consenso acerca da melhor ferramenta a ser utilizada. Adicionalmente, os critérios divergem entre elas, de modo que um mesmo estudo pode ser classificado de maneira diferente a depender da ferramenta utilizada.

Diversos estudos têm avaliado as diferentes ferramentas de acesso ao risco de viés dos estudos observacionais. Nesse sentido, West e cols.⁴⁴ identificaram 12 ferramentas utilizadas em estudos observacionais, e seis destas eram para utilização em revisões sistemáticas. Deeks e cols.⁴⁵ identificaram as 14 “melhores” ferramentas de 182 e recomendaram seis para utilização no contexto de revisões sistemáticas. Dessas duas

evidências, apenas três ferramentas em comum foram reportadas pelos estudos. Ainda, essas três ferramentas em comum podem ser utilizadas para diferentes desenhos de estudo e apenas duas foram desenvolvidas para serem utilizadas no contexto de revisões sistemáticas.

Sanderson e cols.⁴⁶ avaliaram 86 ferramentas específicas para avaliação de risco de viés em estudos observacionais. Assim como nas ferramentas para risco de viés de ensaios clínicos randomizados, essas variaram em relação aos métodos de desenvolvimento e propósito de utilização: 15% foram desenvolvidas especificamente para serem utilizadas em revisões sistemáticas; 36% foram desenvolvidas para contexto geral de tipos de estudo e 34% foram desenvolvidas para serem utilizadas apenas em contextos específicos. Os autores concluíram a avaliação não recomendando nenhuma ferramenta específica, porém identificaram limitações importantes nas ferramentas avaliadas. Assim, sugeriram evitar as que continham poucos componentes ou domínios, as que não eram específicas para o desenho do estudo, as que não descreviam os métodos utilizados para o desenvolvimento da ferramenta e as que não tinham sido validadas. Adicionalmente, identificaram maior limitação nas ferramentas que propunham pontuações, em relação aos *checklists* mais simples, nos quais os principais domínios são avaliados de forma individual.

A Colaboração Cochrane destaca duas ferramentas que podem ser utilizadas para avaliação do risco de viés dos estudos individuais incluídos em revisões sistemáticas: Downs and Black⁴⁷ e Newcastle Ottawa Scale⁴⁸. Além disso, a própria tabela de risco de viés da Colaboração Cochrane, elaborada para avaliação da qualidade metodológica e risco de viés dos ensaios clínicos randomizados pode ser utilizada para os estudos de coorte prospectivos.

O Newcastle Ottawa Scale é um questionário para ser utilizado especificamente no contexto de revisão sistemática e contém questões separadas para estudos de coorte e caso-controle (Anexo C). Contempla os seguintes domínios de avaliação do risco de viés: seleção dos pacientes (generalização e aplicabilidade), comparabilidade dos grupos no estudo, métodos para avaliação dos desfechos (estudos de coorte), comprovação da exposição (caso-controle) e seguimento adequado.

O *checklist* de Downs and Black foi desenvolvido e validado para avaliação da qualidade metodológica de estudos randomizados e observacionais. Assim, algumas questões não são aplicáveis às revisões sistemáticas de estudos observacionais e a Colaboração Cochrane sugere que o instrumento seja adaptado para diferentes contextos. A ferramenta é composta por 27 questões que avaliam a qualidade metodológica nos seguintes domínios:

- Relato: composto por nove questões que avaliam se a informação reportada na publicação foi suficiente para permitir que o leitor faça uma avaliação livre de viés sobre os resultados do estudo;

- Validade externa: três questões que avaliam o quão os achados do estudo são generalizáveis para a população com as mesmas condições daquela incluída no estudo;
- Viés: composto por sete questões que avaliam a presença de viés de aferição diferencial que ocorre quando os métodos de mensuração da exposição ou desfecho são sistematicamente diferentes entre os grupos;
- Variável de confusão/viés de seleção: seis questões que avaliam a presença de viés de seleção dos participantes do estudo, que pode levar à introdução desbalanceada de variáveis de confusão;
- Poder: uma questão que avalia a probabilidade dos achados do estudo ser atribuída ao acaso.

As respostas são pontuadas com o valor “1” (quando o critério que caracteriza qualidade estiver presente) ou “0” (quando o critério que caracteriza qualidade estiver ausente), exceto em uma questão referente ao domínio “relato” na qual são permitidas três respostas (pontuação de 0 a 2). Assim, estudos de melhor qualidade metodológica atingem maior pontuação (Anexo D).

Diante da ausência de consenso sobre a melhor ferramenta para avaliação do risco de viés, recomendamos que os revisores avaliem o risco de viés em cada domínio onde a presença de viés possa distorcer sistematicamente a magnitude e/ou direção da medida de associação. Ademais, é importante reportar as informações que subsidiaram o julgamento para indicar a presença ou ausência do risco de viés. Por exemplo: em um estudo de coorte onde nem todos os pacientes elegíveis tenham sido incluídos consecutivamente, o risco para viés de seleção é alto. Nesse exemplo, é mais recomendado indicar que existe alto risco de viés de seleção e a justificativa. Assim, se o revisor optar pelo uso de um *checklist* como a Downs and Black, que fornece um escore final, recomenda-se que na apresentação do risco de viés dos estudos incluídos o revisor exponha também a justificativa para a pontuação de cada questão/domínio, de forma que o leitor possa compreender a razão para a tomada de decisão. Os Anexos C e D apresentam as traduções simples (não validada) do Newcastle Ottawa Scale e o *checklist* Downs and Black adaptado, respectivamente, acompanhadas dos instrumentos no idioma original.

3.3.1.4 Extração dos dados quantitativos

Os dados quantitativos são aqueles que permitirão o cálculo da metanálise, se aplicável, ou a apresentação sumária das medidas de associação entre fator de exposição e o desfecho.

Em estudos observacionais, duas medidas de associação podem ser apresentadas pelos estudos: a medida de associação não ajustada e a medida de associação ajustada. Conforme discutido em maiores detalhes nas seções subsequentes, a medida de associação não ajustada é análoga à medida de efeito relatada nos ensaios

clínicos randomizados. Seria a medida bruta da associação. A medida ajustada é uma estimativa de efeito que foi estatisticamente corrigida pelos fatores de confusão considerados no estudo.

É importante coletar essas duas medidas, quando disponíveis, pois, conforme será discutido, várias são as abordagens utilizadas para o cálculo da metanálise.

DIRETRIZ: EXTRAÇÃO DE DADOS

- 1) Elaborar ficha padrão com campos para coleta de dados descritivos (características e risco de viés dos estudos incluídos) e dados quantitativos (medidas de associações ajustada e não ajustada).
- 2) Extrair os dados de cada estudo incluído na revisão sistemática.
- 3) A extração dos dados deve ser realizada por dupla de revisores, de forma independente.
- 4) Resolver as discordâncias por consenso ou por um terceiro revisor.

3.4 Resultados

Independente do número de artigos elegíveis para a revisão sistemática, nesta etapa todos os dados de interesse, contidos na ficha de coleta de dados, devem estar extraídos de cada estudo. Esses dados podem ser compilados em forma de tabelas e, se aplicável, podem ser utilizados para cálculo da medida sumária da metanálise.

3.4.1 Apresentação dos dados descritivos

As características clínicas de base da população incluída nos estudos primários de coorte e caso-controle, que foram definidas na ficha clínica, devem ser apresentadas para cada estudo de maneira descritiva, em forma de tabela. Além disso, outras características metodológicas que possam ser relevantes para a interpretação dos resultados também devem ser apresentadas.

Dependendo do periódico-alvo de publicação da revisão sistemática, o número de tabelas pode ser limitado, de forma que características da população e características metodológicas relevantes como delineamento, períodos de seguimento e fatores de confusão considerados tenham que ser condensados em uma mesma tabela.

Outra tabela descritiva importante para avaliação crítica da revisão sistemática é a que apresenta o risco de viés dos estudos incluídos, independente da ferramenta utilizada para avaliação.

O fluxo de seleção dos artigos é outro resultado descritivo. Permite ao leitor uma ideia da abrangência da estratégia de busca e aumenta a validade interna da revisão. Um exemplo de como apresentar esse fluxo encontra-se no Anexo B.

3.4.2 Sumário das medidas de associação nos estudos

Ao contrário da metanálise de ensaios clínicos randomizados, o principal propósito da metanálise de estudos observacionais não é derivar uma única medida de associação, mas investigar as razões das diferenças nas medidas de associação entre os estudos e encontrar padrões de associação entre eles.

Em uma mesma revisão sistemática, o cálculo da metanálise pode ser possível e adequado para alguns desfechos e inadequado para outros. Conforme definido anteriormente, metanálise é o uso de métodos estatísticos para combinar os resultados dos estudos individuais. Entretanto, a combinação dos resultados não corresponde simplesmente à soma das populações dos estudos, como se todos fizessem parte de um mesmo grande estudo. Simplificadamente são consideradas as médias ponderadas das medidas de associação dos estudos incluídos para calcular a metanálise. O tamanho da amostra e o número de eventos são os determinantes dos pesos dos estudos na metanálise. Assim, a medida sumária da metanálise leva em conta os intervalos de confiança e a força da associação desses estudos individuais. A relação entre peso e tamanho de amostra é complexa e vai depender do método estatístico escolhido para realizar a metanálise.

A medida de associação sumária é calculada a partir da identificação da natureza do dado e definição da medida de associação (por exemplo, *odds ratio* ou risco relativo). Paralelamente, define-se o modelo de análise da metanálise e o método estatístico aplicável ao tipo do modelo.

3.4.2.1 Cálculo da metanálise de estudos observacionais

Nos estudos observacionais comparativos, a presença de fatores de confusão representa o principal limitante. Fatores de confusão são variáveis observadas ou não observadas que estão relacionadas com a exposição de interesse e com o desfecho e, dessa forma, podem influenciar a magnitude e direção da medida de associação entre as variáveis.

Análises estatísticas permitem controlar a influência dos fatores de confusão e de interação em nível de estudo primário. O desafio é que, não necessariamente, todos os fatores importantes são conhecidos e/ou considerados nas análises. Essa é a principal razão pela qual os ensaios clínicos randomizados são hierarquicamente superiores aos estudos observacionais: a randomização adequada (considerando um tamanho amostral adequado) permite criar grupos verdadeiramente comparáveis, tanto para os fatores conhecidos quanto para os fatores desconhecidos, de modo a atribuir a diferença nos achados exclusivamente à intervenção.

No relato dos resultados dos estudos observacionais primários, se os autores não considerarem ajustar a associação entre o fator de exposição e o desfecho por

nenhuma variável de confusão, ele irá, então, apresentar a estimativa não ajustada dos resultados. Os achados não ajustados representam a relação bivariada entre a variável independente e a variável dependente que não foi controlada por fatores de confusão. São análogos aos resultados gerados por ensaios clínicos randomizados, que atribuem as diferenças entre os achados exclusivamente à exposição.

Os autores dos estudos individuais podem controlar a influência das variáveis de confusão no resultado por meio de modelos de regressão múltipla, onde cada fator de confusão é inserido no modelo simultaneamente com a variável independente principal (fator de exposição). Nesses casos, o resultado bruto sofrerá um “desconto” da influência desses fatores de confusão e resultará em uma medida “ajustada”.

Outra opção de ajuste estatístico para variáveis de confusão nos estudos primários é o escore de propensão. A técnica utiliza as potenciais variáveis de confusão para estimar a probabilidade de cada indivíduo participante apresentar o desfecho de interesse. Essa probabilidade é então “descontada” na análise para gerar a medida de associação. Então, a medida de associação estará livre da influência das variáveis de confusão. Existe ainda, o escore de propensão pareado, onde os grupos exposto e não exposto são pareados de acordo com escore de propensão (a probabilidade de apresentar o desfecho com base nas variáveis de confusão), gerando assim, grupos comparáveis (com distribuição das variáveis de confusão balanceadas).

Adicionalmente, as análises estratificadas podem também ser utilizadas onde a população do estudo é dividida em estratos, que representam as variáveis de confusão ou de interação, e as análises são conduzidas separadamente por estrato.

Considerando os diferentes tipos de estudos observacionais (coortes ou caso-controle) e suas distintas fontes de vieses, a primeira recomendação é que não sejam consideradas as estimativas dos diferentes delineamentos de estudo na mesma metanálise⁴¹. Se os critérios de inclusão contemplarem estudos observacionais de ambos os delineamentos, duas estimativas de metanálise deverão ser apresentadas, uma para os estudos de coorte e outra para os estudos de caso-controle, se possível. Obviamente, deverá ser discutido o quão as duas estimativas convergem ou divergem quanto a tendências ou magnitude de associação. Quando se considera a influência dos fatores de confusão nos resultados dos estudos, não há um consenso documentado de como sintetizar os achados ajustados e não ajustados.

É preciso lembrar que mesmo a metanálise de ensaios clínicos randomizados está sujeita à influência da heterogeneidade entre os diferentes estudos incluídos, mesmo que essa heterogeneidade não exista dentro de cada estudo individual. No caso de estudos observacionais, as variáveis de confusão que foram consideradas nos ajustes estatísticos também variam de estudo para estudo e essa variabilidade é uma fonte adicional de heterogeneidade entre os estudos. Por exemplo, um estudo que queira

estimar a associação entre ingestão de gordura monoinsaturada e a ocorrência de infarto do miocárdio em coortes de pessoas residentes em países do mediterrâneo pode ter corrigido o efeito da ingestão por diferenças na idade, sexo e índice de massa corporal. Outro estudo, avaliando a mesma questão de pesquisa, pode ter considerado as mesmas variáveis, porém incluído também a influência do histórico de doença cardiovascular prévio no modelo de regressão múltipla. Nesse sentido, se o revisor considerar que, de fato, o histórico de doença cardiovascular é uma variável de confusão que pode influenciar nos achados, a questão é: como a diferença no número de variáveis e das variáveis em si (consideradas nos modelos dos estudos individuais) é ponderada na metanálise?

Uma maneira simplista de resolver essa questão é não considerar a heterogeneidade devido ao ajuste por covariáveis e incluir na metanálise apenas os resultados não ajustados. Esse tipo de análise fornece resultados de difícil interpretação, pois é o tipo de análise que se utiliza na metanálise de ensaios clínicos randomizados, que tem na randomização o meio de “eliminar” o efeito das variáveis de confusão e outros vieses, o que não é o caso dos estudos observacionais.

Alguns revisores preferem considerar na metanálise as medidas de associação ajustadas para um conjunto de variáveis de confusão pré-determinadas e incluir no cálculo apenas os estudos que ajustaram seus modelos para esse conjunto de covariáveis^{49,50}. O limitante dessa abordagem é que, conforme descrito anteriormente, nem todos os estudos ajustam seus modelos de análises considerando as mesmas variáveis. Assim, essa abordagem implica em restringir o número de evidências disponíveis para responder a questão de pesquisa, contrariando o princípio básico da revisão sistemática de apresentar toda a evidência disponível.

Outra opção é incluir os resultados ajustados e não ajustados na mesma metanálise⁵¹. Essa abordagem permite incluir toda a evidência recuperada, mas não leva em consideração a heterogeneidade devido aos ajustes nas análises. O limitante dessa abordagem é que análises não ajustadas geralmente produzem medidas de associação de maior magnitude e, dessa forma, a inclusão dessas medidas pode superestimar a medida de associação entre exposição e desfecho. Assim, na ausência de uniformidade dos relatos entre os estudos, uma maneira alternativa e mais adequada de análise é calcular separadamente a metanálise das estimativas ajustadas e das estimativas não ajustadas. Muitas vezes elas tendem para a mesma conclusão, apresentando diferenças na magnitude da associação, que pode ser discutida com base no acima exposto. As dificuldades de interpretação emergem quando a metanálise das estimativas ajustadas e não ajustadas não sustentam a mesma conclusão.

Independente da abordagem metanalítica utilizada, essa diferença nos fatores de confusão considerados nos estudos primários muito provavelmente repercutirá na significância e magnitude da heterogeneidade da metanálise. É por isso que a exploração da heterogeneidade é uma questão essencial em metanálise de estudos observacionais (seção 3.4.5).

3.4.2.2 Modelos de análise

Uma vez identificada a natureza do dado e estabelecida qual será a medida de associação calculada, devemos definir o modelo de análise da metanálise. Dentre esses modelos, pode-se optar pelo modelo de efeito fixo ou de efeitos randômicos, conforme descritos a seguir.

3.4.2.2.1 Modelo de efeito fixo

O modelo de efeito fixo de metanálise baseia-se no pressuposto matemático de que existe um único verdadeiro efeito comum entre os estudos e que as diferenças entre as medidas de associação dos estudos são atribuídas meramente ao acaso. Neste caso, se todos os estudos fossem infinitamente grandes, obteriam as mesmas medidas de associação. O modelo de efeito fixo requer duas suposições. A primeira é assumir que cada estudo mede o mesmo parâmetro. A segunda é que não existem variações nas fontes de população.

Assim, como os estudos observacionais são heterogêneos por sua natureza, o modelo de efeito fixo não é adequado para as análises.

3.4.2.2.2 Modelo de efeitos randômicos

O modelo de efeitos randômicos de metanálise assume que a verdadeira medida de associação pode ser diferente entre os estudos primários, ou seja, não existe um único valor que estima a medida sumária da metanálise e sim uma distribuição de números.

Os modelos de efeitos randômicos e efeito fixo proverão resultados idênticos quando não há heterogeneidade entre os estudos. Caso exista heterogeneidade entre os estudos, o modelo de efeitos randômicos proverá uma medida de associação com menor precisão (isto é, com intervalo de confiança mais amplo), sendo uma abordagem mais conservadora e mais indicada para os estudos observacionais.

3.4.2.3 Métodos estatísticos

Diferentes métodos estatísticos são utilizados para ponderar as medidas de associação entre os estudos incluídos na revisão. Assim como em metanálise de ensaios clínicos randomizados, as mesmas abordagens podem ser utilizadas, conforme descritas abaixo.

3.4.2.3.1 Método do inverso da variância

O método do inverso da variância pode ser usado para combinar dados dicotômicos ou contínuos e para modelos de efeito fixo e efeitos randômicos. Tem alta aplicabilidade, pois combina qualquer medida de associação que tenha reportado erro padrão ou intervalo de confiança.

O conceito matemático de variância é o quadrado do erro padrão. Por sua vez, erro padrão descreve a extensão em que a estimativa pode estar errada devido ao erro aleatório. Estes dois conceitos são intercambiáveis em termos de informação, mas com valores numéricos diferentes.

Pelo método do inverso da variância, se assume que a variância é inversamente proporcional à importância do estudo, ou seja, quanto menor a variância, mais peso será atribuído a este estudo.

3.4.2.3.2 Método Mantel-Haenszel

O método de Mantel-Haenszel é utilizado para modelos de efeito fixo e efeitos randômicos, mas somente quando os dados forem dicotômicos. Este método é preferível ao método do inverso da variância quando os dados dos estudos são escassos em termos de eventos e pequenos tamanhos de amostra, pois utiliza um modelo diferente de atribuição de pesos que o inverso da variância. Nas demais situações, este método resulta em estimativas similares ao método do inverso da variância.

3.4.2.3.3 Método de Peto

O método de Peto é utilizado somente para dados dicotômicos em que o *odds ratio* foi reportado como medida de associação e se o modelo eleito foi o modelo de efeito fixo.

É um método alternativo ao método de Mantel-Haenszel, preferível quando os dois grupos têm grosseiramente o mesmo número de participantes e a magnitude da medida de associação é pequena, mas significativa. Ainda, é uma melhor abordagem na estimativa de *odds ratios* quando a ocorrência de eventos é igual à zero em um dos grupos.

3.4.2.4 Programas para cálculo de metanálise

Diversos programas estatísticos permitem esse cálculo (R, SAS, PLINK). Nesses casos, o método do inverso da variância é utilizado e há necessidade de conhecer os comandos de cada programa.

Para o método de Mantel-Haenszel e Peto, são necessários os dados da tabela de contingência 2 X 2. Nesse sentido, existem métodos que permitem reconstruir a tabela, através de imputação de dados^{52, 53}.

Uma série de programas estatísticos, gratuitos e comerciais, pode ser utilizada para calcular a metanálise. Uma opção gratuita é o programa estatístico “R”⁵⁴, que possui alguns pacotes básicos instalados e outros que devem ser incorporados separadamente, como o metafor, metacor e metaMA. Outras opções, também gratuitas, são o Meta-

Analysis Version 5.3⁵⁵ e o Metaanalysis with Interactive Explanations MIX 2.0⁵⁶. Algumas opções de programas comerciais são o Comprehensive Meta-Analysis CMA⁵⁷, MetaWin⁵⁸, WEasyMa⁵⁹, que também possuem opções demo gratuitas, o SAS⁶⁰, STATA⁶¹ e o WinBUGS⁶².

O Review Manager⁶³ é uma das ferramentas mais utilizadas para cálculo de metanálises de estudos randomizados. Pode ser utilizado para cálculo de metanálise de estudos de coorte quando os estudos incluídos reportarem o número de eventos no grupo exposto e não exposto pelos métodos de Mantel-Haenszel ou Peto, neste caso, gera uma medida de associação não ajustada. O Review Manager também permite a sumarização das medidas de associação reportadas pelos estudos, por exemplo, do risco relativo ou *odds ratio* transformado em escala logarítmica (para que a escala numérica se torne simétrica).

Portanto, para o cálculo da metanálise de uma medida de associação em escala logarítmica no Review Manager deve-se selecionar como tipo de dado “inverso da variância genérica”. Assim, é possível inserir dados da medida de associação de escolha e seu erro padrão. O método de análise é o inverso da variância, que permite calcular a metanálise para modelos de efeito fixo ou efeitos randômicos. O gráfico da metanálise gerada também é apresentado em escala logarítmica, assim os intervalos de confiança parecem simétricos. O Apêndice F discorre sobre cálculo e interpretação das medidas de associação.

3.4.3 Apresentação dos resultados das metanálises

A(s) medida(s) de associação da(s) metanálise(s) pode(m) ser apresentada(s) em forma de tabelas ou gráficos. O gráfico de floresta ou “*forest plot*” é a forma gráfica de apresentação, geralmente mais usual e informativa, que exhibe a medida de associação e respectivo intervalo de confiança para os estudos individuais e para a metanálise, além de permitir uma inspeção visual qualitativa sobre as tendências dos estudos individuais e metanálise. A interpretação de um gráfico de floresta encontra-se no Apêndice G.

3.4.4 Risco de viés de publicação

Além dos diversos tipos de vieses (por ex.: de seleção, de aferição e as variáveis de confusão) que podem afetar as medidas de associação dos estudos observacionais e, portanto, das metanálises dos mesmos, em revisões sistemáticas ainda soma-se o risco de viés de publicação. Este viés é caracterizado pela não inclusão da totalidade da evidência disponível, e é geralmente introduzido quando o processo de busca não é abrangente ou quando os estudos não publicam os resultados de todos os desfechos originalmente planejados (geralmente aqueles negativos).. A presença do viés de publicação pode levar à distorção da medida de associação, geralmente na direção da superestimação da associação⁶⁴. Isto ocorre devido à maior probabilidade de publicação de estudos cujos resultados sejam positivos, ou seja, que apresentam

associação entre exposição e desfecho (viés de citação). Os principais tipos de vieses relacionados às revisões sistemáticas estão resumidos na Quadro 7 a seguir, e todos eles podem introduzir viés de publicação na revisão sistemática.

Quadro 7 – Principais vieses de publicação

Tipo de viés	Definição
Viés de tempo para publicação	Rápida ou demorada publicação dos resultados da pesquisa, dependendo da natureza e direção dos resultados.
Viés de múltipla publicação	Resultados da pesquisa publicados em várias fontes, dependendo da natureza e direção dos resultados, que podem ser computados mais de uma vez na metanálise.
Viés do local de publicação	A publicação dos resultados em revistas de fácil acesso ou níveis de indexação em base de dados padrão, dependendo da natureza e direção dos resultados.
Viés de citação	A citação ou não citação dos resultados da pesquisa, dependendo da natureza e direção dos resultados, com tendência a menor citação de estudos com resultados negativos.
Viés de idioma	A publicação dos resultados da pesquisa em uma língua específica, dependendo da natureza e direção dos resultados ou exclusão de estudos publicados em outro idioma, no processo de seleção dos artigos.
Viés de publicação seletiva de desfecho	O relato de resultados para alguns desfechos e não de outros, dependendo da natureza e direção dos achados.

Fonte: elaboração própria.

Estudos que apresentam uma associação entre exposição e desfecho tendem a ser mais publicados que estudos onde existe ausência de associação⁶⁵. Há muitos fatores que predispoem a este problema. Os autores podem desistir de escrever o artigo por acharem que não há interesse nos resultados ou os editores das revistas podem não demonstrar interesse em publicar o artigo devido ao menor interesse por parte dos leitores. Por isso a importância de pesquisar fontes de dados não publicados e reunir a totalidade da evidência.

Alguns testes estatísticos permitem avaliar o viés de publicação. Todas possuem vantagens e limitações que devem ser consideradas na hora da escolha. Ainda, as evidências que comparam as diversas técnicas são limitadas e necessitam ser mais bem investigadas. Abaixo, uma breve descrição dos métodos mais conhecidos de acesso ao viés de publicação.

3.4.4.1 Gráfico de funil

O gráfico de funil ou *funnel plot* é o método mais utilizado para avaliar a presença de vieses de publicação e consiste em um *scatter plot* onde são plotados no eixo horizontal a medida de associação e no vertical alguma medida que represente o poder estatístico e a precisão dos resultados dos estudos individuais.

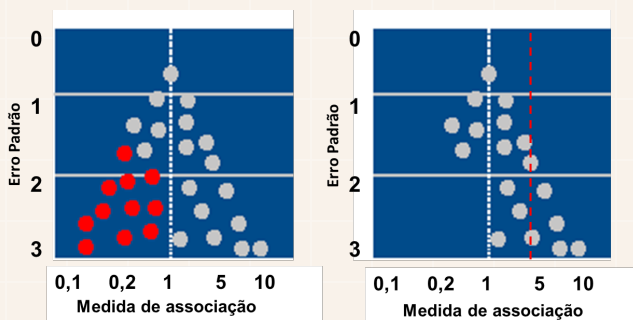
O nome “gráfico de funil” vem do fato de que na ausência de viés de publicação os pontos plotados (que representam as medidas de associação dos estudos individuais) são simetricamente distribuídos em torno da linha de verdadeira associação formando um padrão semelhante a um funil invertido (Figura 3).

A precisão da medida de associação aumenta à medida que o poder estatístico aumenta. Geralmente, estudos mais precisos e com maior poder estatístico são também os com maiores amostras. Na base, estariam os estudos menos precisos, que podem apresentar variações nas tendências da medida de associação por mero acaso. À medida que estudos maiores são publicados, a medida de associação tende a permanecer a mesma, cada vez mais precisa, configurando o vértice do funil.

Na presença de viés, por exemplo, devido a não publicação de estudos pequenos nos quais a medida de associação foi menor que o verdadeiro efeito (se encontram entre o verdadeiro efeito e a linha de nulidade) ou não houve associação entre exposição de desfecho ou a associação se demonstrou inversa à hipótese inicial do estudo, os pontos são assimetricamente distribuídos e, neste exemplo, haverá ausência de pontos na base esquerda da pirâmide.

Na maioria das vezes, a variável plotada no eixo vertical é o erro padrão. Outras variáveis como variância, inverso da variância e inverso do erro padrão também podem ser plotados no eixo vertical. No entanto, não se recomenda plotar o tamanho da amostra no eixo vertical, pois o poder estatístico e a precisão de um estudo são determinados por outros fatores além do tamanho de amostra, como o número de participantes que apresentam os desfechos de interesse (para desfechos dicotômicos) e o desvio padrão (para desfechos contínuos). Por exemplo, um estudo com 100.000 participantes e 10 eventos tem menor poder estatístico e fornecerá uma medida de associação menos precisa que um estudo com 1.000 participantes e 100 eventos. Assim, estudos com tamanhos de amostras muito diferentes podem apresentar o mesmo erro padrão e precisão (e vice-versa)⁶⁶. A Figura 3 ilustra a dispersão do gráfico de funil, em diferentes situações.

Figura 3 – Gráfico de funil hipotético



Fonte: (Sterne e cols., 2001, adaptado)⁶⁶.

O eixo vertical mostra o erro padrão de cada estudo, ou seja, quanto menor o erro padrão, maior a precisão do estudo (e geralmente maior tamanho de amostra). No eixo horizontal, está representando a medida de associação de cada estudo. À esquerda, ausência de sinais sugestivos de viés de publicação. Os círculos vermelhos na figura da esquerda representam os estudos nos quais a medida de associação foi menor que o verdadeiro efeito, nos quais não houve associação entre exposição e desfecho ou a associação se demonstrou inversa à hipótese inicial do estudo. À direita, um gráfico com dispersão que representa viés de publicação onde esses estudos não foram publicados (ou não foram localizados para inclusão na revisão sistemática). Neste caso, a metanálise obtida é representada pela linha vermelha e sugere falsamente associação positiva.

Uma limitação do gráfico de funil para detecção de viés de publicação é que algumas medidas de associação, como *odds ratio* e diferenças de médias padronizadas, estão naturalmente correlacionadas com seus erros padrão e podem produzir assimetria no gráfico de maneira que não necessariamente represente viés de publicação. Assim, é importante deixar claro que a assimetria no gráfico de funil não quer dizer obrigatoriamente que há viés de publicação.

Outras razões que podem levar à assimetria no gráfico de funil devem ser investigadas pelos revisores, incluindo heterogeneidade entre os estudos ou o acaso. Além disso, a qualidade metodológica dos estudos incluídos pode ser outro motivo de assimetria. Isso porque estudos pequenos tendem a ser delineados e conduzidos com menor rigor metodológico e são mais suscetíveis a vieses⁶⁷. Esses estudos tendem também a produzir associações positivas espúrias⁶⁸. Esse fenômeno pode ser denominado “efeito de pequenos estudos”. Além disso, na presença desse fenômeno, alguns estudos que apresentariam resultados “negativos” caso fossem delineados e conduzidos adequadamente, podem ter apresentado resultados “positivos”, causando assimetria no gráfico⁴¹.

Alguns autores têm argumentado que a interpretação visual do gráfico é muito subjetiva e tem sua aplicabilidade limitada⁶⁹. Além disso, quando existirem poucos estudos incluídos na metanálise, a inspeção visual do gráfico pode ficar comprometida. Alguns testes estatísticos foram desenvolvidos no intuito de avaliar as causas para a assimetria do gráfico de funil. Os testes para assimetria de funil estão descritos a seguir.

3.4.4.2 Teste para assimetria do gráfico de funil

Os testes para avaliar a causa da assimetria do gráfico de funil avaliam se a relação encontrada entre a medida de associação e o erro padrão (ou outra variável que reflita a precisão e tamanho da amostra) ocorreu pelo acaso, ou seja, podem detectar se a causa da assimetria é o “efeito de estudos pequenos” (estudos menores que, geralmente, são conduzidos com menos rigor metodológico e tendem a apresentar maiores medidas de associação)⁴¹. Vários testes foram desenvolvidos com esse

propósito, no entanto, o teste proposto por Egger e cols. tem sido o mais utilizado para avaliar a assimetria do funil⁷⁰.

O teste de regressão de Egger⁷⁰ é uma abordagem que realiza uma regressão linear da estimativa do efeito da intervenção pelo seu erro padrão, ponderada pelo inverso da variância da medida de associação. Essa ponderação é importante, pois assegura que a estimativa de regressão não seja dominada por estudos pequenos.

Para desfechos contínuos com diferenças de média como medida de associação, o teste de regressão de Egger deve ser preferido aos outros testes. Entretanto, quando o desfecho é dicotômico e a medida de associação da metanálise é o *odds ratio*, a taxa de erro tipo I (proporção de resultados falso-positivos) é alta. Ou seja, mesmo na ausência de “efeito de pequenos estudos”, o teste de Egger pode indicar que a associação entre as características metodológicas e a magnitude da medida de efeito não ocorreu ao acaso (apresentou valor de *p* muito pequeno). Alguns outros testes diminuem a correlação entre o logaritmo do *odds ratio* e sua variância, diminuindo estas altas taxas de falso-positivos⁷¹. Entretanto, estes testes aumentam resultados falso-positivos na presença de heterogeneidade significativa.

Quando o teste para assimetria do gráfico de funil indicar “efeito de pequenos estudos”, o viés de publicação deve ser considerado como apenas uma dentre outras possíveis explicações. É importante salientar que os testes para avaliar a assimetria do gráfico de funil devem ser utilizados apenas quando houver no mínimo dez estudos incluídos na metanálise. Quanto menos estudos são incluídos o poder dos testes é ainda menor⁴¹.

3.4.4.3 Método “trim” e “fill”

O método “trim” e “fill”⁷² é um método não paramétrico que objetiva identificar e corrigir a assimetria do gráfico de funil que é proveniente somente do viés de publicação. É uma abordagem baseada no princípio da simetria, que estima o número de estudos que deveriam estar presentes no gráfico.

A base do método é excluir (cortar, do inglês “trim”) os estudos pequenos e causadores da assimetria e usar o funil “recortado” para estimar o verdadeiro centro do funil. Em seguida, substitui-se os estudos omitidos e seus equivalentes, do outro lado, ao redor do centro, preenchendo (completar, do inglês “fill”) o funil. Desta forma, estima o número de estudos faltantes (não publicados) e uma medida de associação ajustada, que é derivada do cálculo de uma nova metanálise que inclui as medidas de associação destes estudos não publicados.

As limitações inerentes a esta abordagem estão pautadas na suposição do modelo, que assume que a assimetria do funil representa o viés de publicação. Se a assimetria for

decorrente de um erro matemático do modelo adotado ou devido à heterogeneidade significativa, esta medida de associação não estará correta, pois não haveria pequenos estudos com mesma magnitude, mas tendências opostas de efeito para contribuir com a estimativa global da metanálise.

3.4.5 Heterogeneidade

Dada a natureza dos estudos observacionais, é esperada a presença de heterogeneidade. Mais importante que derivar uma medida única de associação, o objetivo prioritário da metanálise é explorar a heterogeneidade. Esse tema é tão relevante que até se a heterogeneidade quantificada não for significativa, a exploração do efeito das possíveis variáveis de confusão ou de interação identificadas deve ser realizada⁷³.

A heterogeneidade pode ocorrer devido aos critérios de elegibilidade da amostra incluída, dos métodos para determinar e quantificar o fator de exposição, nas variáveis de confusão controladas nas análises, nos métodos para diagnosticar ou definir os desfechos, etc.⁷⁴ Os estudos também podem ser heterogêneos em relação às suas características metodológicas e estar mais ou menos suscetíveis a vieses.

Assim como em metanálises de ensaios clínicos randomizados, quando a heterogeneidade é muito elevada e não pode ser explicada por nenhuma análise ou de subgrupo, não se recomenda a realização da metanálise, e as medidas de associação entre exposição e desfecho dos estudos não podem ser apresentados de forma agregada.

3.4.5.1 Quantificação de heterogeneidade

Alguns testes estatísticos conseguem mensurar o grau de heterogeneidade entre os estudos. O teste do chi-quadrado (χ^2 ou X^2) mensura o quanto a diferença de magnitude entre as medidas de associação são atribuídas somente ao acaso. Porém, este teste apresenta algumas limitações expressivas, como não fazer distinção entre estudos com tamanho de amostra insuficiente daqueles com maior tamanho de amostra. Além disso, o teste de χ^2 tem baixo poder estatístico. Um valor de p pequeno para o teste de χ^2 (geralmente convencionado em $<0,10$ devido ao baixo poder) pode indicar a presença de heterogeneidade, ou seja, a variabilidade entre as medidas de associação não existe apenas devido ao acaso. Por outro lado, um teste não significativo não exclui totalmente a presença de heterogeneidade.

Um teste estatístico de maior abrangência é o teste de inconsistência de Higgins ou I^2 .⁷⁵ O teste de I^2 leva em conta o resultado do χ^2 além da magnitude e direção das medidas de associação. Alguns autores alertam que sempre existirá alguma heterogeneidade em metanálise e o teste de I^2 descreve o quanto a heterogeneidade

impacta na medida de associação. A interpretação do I^2 deve levar em conta as possíveis fontes de heterogeneidade clínica e metodológica, mas simplificadaamente pode-se interpretar o I^2 da seguinte forma:

- 0% a 25% – heterogeneidade pode não ser importante;
- 25% – 50% – heterogeneidade moderada;
- > 50% – heterogeneidade alta.

3.4.5.2 Explorando a heterogeneidade entre os estudos

É possível explorar as causas de heterogeneidade da metanálise por meio de meta-regressão ou análises de subgrupo. Essas análises objetivam verificar se determinadas características podem modificar a magnitude da associação entre exposição e desfecho (fator de interação). Além disso, podem verificar se determinadas variáveis são responsáveis pela presença da associação (variável de confusão).

3.4.5.2.1 Meta-regressão

Em metanálise de estudos observacionais, considerando o delineamento de estudo observacional e suas limitações, a meta-regressão é a abordagem mais adequada para explorar as fontes de heterogeneidade entre os estudos.

A meta-regressão é, em essência, uma análise de regressão onde um desfecho pode ser predito com base em uma ou mais variáveis explanatórias. Na meta-regressão a variável desfecho é a medida de associação (por exemplo, o risco relativo ou *odds ratio*) e as variáveis explanatórias são características dos estudos que possam influenciar na magnitude e/ou direção da medida de associação. Essas variáveis explanatórias são denominadas covariáveis ou ‘potenciais modificadoras de efeito’, ou seja, são potenciais variáveis de confusão ou interação.

Se a medida de associação é uma razão (por exemplo, risco relativo) deve-se utilizar o logaritmo da razão da medida de associação no modelo de meta-regressão. O coeficiente de regressão gerado pela análise da meta-regressão descreve o quanto a variável desfecho (medida de associação) é modificada por unidade de incremento da variável explanatória.

Especificamente no contexto dos estudos observacionais, uma das variáveis que pode ser considerada na meta-regressão é o número de fatores de confusão ajustados nos modelos de regressão múltipla nos estudos individuais. Se a meta-regressão apresentar influência do número de fatores de confusão na magnitude da heterogeneidade encontrada na metanálise, pode-se discutir a influência da seleção e a quantidade desses fatores de confusão nos modelos de regressão múltipla nos diferentes estudos. Por exemplo, Horta e col.⁷⁶ definiram que idade, raça e nível socioeconômico poderiam influenciar a relação entre tabagismo materno e desmame precoce. Na meta

regressão, incluíram, então, a covariável categórica “modelo não ajustado”, “modelo parcialmente ajustado” ou “modelo totalmente ajustado” baseado nessas variáveis. Eles concluíram que o ajuste para essas variáveis de confusão respondiam por 14,4% da heterogeneidade da medida de associação da metanálise.

Outra alternativa para considerar a diferença no número de fatores de confusão utilizados nos estudos individuais é gerar uma covariável dicotômica para a meta-regressão, onde um valor de corte, que representaria o número de fatores de confusão utilizados nos modelos de regressão múltipla, definido de acordo com algum critério estabelecido pelos revisores, permitiria considerar os estudos como “menor que” e “maior que” o valor de corte definido^{77, 78}.

É importante especificar as variáveis que serão incluídas no modelo de meta-regressão antecipadamente com base em um racional para a exploração de cada característica. Análises pré-especificadas diminuem a probabilidade de resultado espúrio.

Assim como nas análises de regressão, o número de variáveis que podem ser controladas simultaneamente na meta-regressão depende do número de estudos incluídos, sendo necessários 10 estudos para cada variável ajustada⁴¹. Essa é a principal limitação dessa abordagem, já que o número de estudos incluídos em metanálises é moderado, na maioria das vezes, o que inviabiliza a exploração da heterogeneidade por todas as covariáveis necessárias.

3.4.5.2.2 Análise de subgrupo (ou análise estratificada)

Essa análise consiste na condução das análises em subgrupos (estratos) de forma a comparar o comportamento da medida de associação nesses estratos. As análises podem ser conduzidas para subgrupos de participantes (por exemplo, pacientes com doença grave *versus* pacientes com doença leve) ou para subgrupos de estudos (por exemplo, estudos conduzidos em diferentes regiões geográficas ou estudos com alto risco de viés *versus* baixo risco de viés).

As análises de subgrupo são conduzidas como forma de explorar a heterogeneidade ou para responder questões específicas a respeito de determinados subgrupos de pacientes, de fatores de exposição ou de estudos com certas características.

Os subgrupos são definidos com base em alguma característica que possa influenciar a medida de associação, geralmente um fator de confusão ou de interação previamente identificado.

Assim como na meta-regressão, é importante especificar as variáveis que definirão os subgrupos previamente e com base em uma justificativa para a exploração de cada característica. Análises pré-especificadas diminuem a probabilidade de resultado

espúrio. Ademais, é importante limitar o número de subgrupos explorados, a fim de minimizar a probabilidade de um achado falso positivo.

As análises de subgrupos apresentam limitações. A principal é considerar apenas uma variável, quando na verdade, pela natureza do estudo, diversas outras variáveis podem estar influenciando na medida de associação. Então, mesmo quando existir associação estatisticamente significativa em um subgrupo enquanto em outro a associação não estiver presente, isso não indica que o fator que estratifica os grupos é responsável por toda a heterogeneidade. Como os subgrupos contêm outras características que podem estar relacionadas à ocorrência ou modificação da associação, é uma visão extremamente simplista comparar apenas a significância estatística das análises de subgrupos.

Análises de subgrupo e meta-regressão devem ser interpretadas com bastante cautela e servem apenas para gerar especulações. Portanto, não devem ser consideradas nas conclusões e recomendações para a prática clínica, embora possam implicar em questões não investigadas, sugerindo a necessidade de novas pesquisas na área. É sempre importante ter uma justificativa para embasar a escolha e o comportamento dos subgrupos ou das análises de meta-regressão.

3.4.6. *Análise de sensibilidade*

O processo para condução de uma revisão sistemática deve ser baseado em critérios objetivos, no entanto, algumas decisões são tomadas de forma arbitrária, muitas vezes por ausência de consenso na literatura para embasar a decisão. Por exemplo, durante a redação do protocolo, pode-se decidir por utilizar determinado método estatístico ou determinado critério de elegibilidade sem se ter certeza de qual é a maneira ideal.

Análise de sensibilidade consiste em repetir a análise para o desfecho primário substituindo variáveis ou testes pré-definidos por decisões alternativas, a fim de demonstrar que os resultados das análises primárias são independentes dessas decisões arbitrárias.

Algumas análises de sensibilidade podem ser definidas na etapa de elaboração do protocolo, contudo, muitas das variáveis adequadas para análise são identificadas durante o processo de revisão, quando particularidades individuais dos estudos são levantadas. Quando as análises de sensibilidade demonstram que os resultados da metanálise não são modificados pelas diferentes decisões que possam ter sido tomadas no processo da revisão, esses resultados podem ser considerados mais confiáveis.

Uma análise de sensibilidade comumente realizada é aquela que exclui o estudo que apresenta uma medida de associação que diverge muito das demais (denominados *outliers*). Esses estudos tendem a distorcer a medida de associação, superestimando

ou subestimando-a. Assim, é possível realizar a metanálise com e sem *outliers*, a fim de verificar sua real influência na magnitude da medida de associação.

Também é importante realizar análise de sensibilidade considerando o risco de viés dos estudos contidos na metanálise e a fim de verificar se o tamanho e/ou a direção da medida de associação são mantidos.

As análises de sensibilidade são, às vezes, confundidas com as análises de subgrupo. Embora algumas análises de sensibilidade consistam em restringir a análise para um subconjunto de estudos, esta análise não determina a medida de associação no conjunto de estudos excluídos da análise. Além disso, análises de sensibilidade propiciam comparações informais sobre diferentes maneiras para se determinar a mesma medida de associação enquanto as análises de subgrupos geram comparações formais sobre as diferentes medidas de associação entre os subgrupos selecionados.

DIRETRIZ: RESULTADOS

- 1)Apresentar tabela com características de base relevantes dos estudos incluídos.
- 2)Apresentar fluxo de seleção dos artigos.
- 3)Determinar os parâmetros estatísticos adequados à revisão.
- 4)Explorar as fontes de heterogeneidade da metanálise.
- 5)Avaliar o viés de publicação da revisão sistemática, se aplicável.

4 Relato e aplicabilidade dos resultados

4.1 Estruturando a discussão e a conclusão

Após obtenção dos resultados, é necessário discuti-los no contexto da própria revisão sistemática e em relação à literatura prévia. O mais importante é interpretar os achados e discutir as fontes de heterogeneidade. Quando a heterogeneidade dos estudos impede a metanálise, os autores da revisão sistemática precisam resumir os resultados e, com base na força de estudos individuais, fazer recomendações na conclusão.

4.1.1 Estruturando a discussão

Uma discussão estruturada ajuda no relato das considerações e implicações da revisão para a prática clínica. Alguns tópicos, baseados no Handbook da Cochrane⁴¹ e na literatura⁷⁹ auxiliam nessa estruturação.

Uma discussão objetiva interpretar os resultados previamente apresentados à luz das evidências atuais e como os novos conhecimentos trazidos com o estudo conduzido podem auxiliar a prática clínica.

Embora a discussão esteja conectada à introdução por meio da hipótese ou questão de pesquisa, não é simplesmente uma repetição ou um rearranjo desta seção. Na verdade, a discussão é um conjunto de ideias que foram construídas desde sua hipótese original, na introdução, até os resultados finais do trabalho, tomando como base os alicerces já construídos por outros estudos e adicionando às evidências os ganhos, concordâncias ou contradições com seus achados.

De uma forma estruturada, inicia-se a discussão com um resumo dos pontos relevantes e fortes dos achados, procurando não repetir os principais resultados de forma quantitativa e limitando-se ao primeiro parágrafo. É importante salientar que na seção dos resultados são apresentados os achados e na discussão são explicados esses achados.

Em seguida, é importante discutir sobre os mecanismos causais e os mecanismos subjacentes à constatação, enriquecendo com os achados concordantes e/ou discordantes de revisões sistemáticas ou estudos importantes publicados previamente.

Por fim, é recomendado apresentar as limitações e os pontos fortes da revisão. Nesta seção, deverá ser avaliada a qualidade dos estudos ou validade interna, a direção e a magnitude dos efeitos. A presença de heterogeneidade é esperada em metanálise de estudos observacionais e sempre será uma limitação inerente. Outra

limitação é ausência de recursos para explorar a heterogeneidade. Por exemplo, às vezes não é possível realizar as análises de subgrupos pré-especificadas pela ausência do relato dos dados desses subgrupos nos estudos primários. Os revisores devem se esforçar para obter os dados não disponíveis, mas nem sempre é fácil obter dados agregados não publicados ou dados individuais dos participantes dos estudos. Também podem ser limitações os potenciais vieses nos estudos primários. Pode ser ressaltada como ponto forte a busca abrangente e baixa probabilidade de vies de publicação (caso os revisores tenham elaborado uma estratégia em todas as potenciais fontes de evidências e não tenham detectado indícios de vies de publicação) ou ainda uma medida de associação robusta e precisa (por exemplo, magnitude do *odds ratio* com intervalo de confiança estreito) que favorece a hipótese de causalidade entre a exposição e desfecho.

No quadro abaixo seguem, resumidamente, as sugestões de estrutura para discussão:

Quadro 8 – Estruturação da discussão da revisão sistemática

- 1 – Resumir os principais resultados.
- 2 – Discutir a abrangência geral e aplicabilidade da evidência.
- 3 – Concordâncias e discordâncias com outros estudos e revisões.
- 4 – Destacar os pontos fortes da revisão.
- 5 – Potenciais vieses e limitações da revisão.

Fonte: (Balsheem e cols., 2011, adaptado)⁸¹.

4.1.2 Estruturando a conclusão

A conclusão pode sumarizar as ideias principais desenvolvidas na discussão e deve ter relação com a resposta ao objetivo da revisão sistemática. Alguns periódicos sugerem que a diretriz *Meta-analysis of observational studies in epidemiology* (MOOSE) seja utilizada para estruturar a discussão (ver item 4.2)². A Cochrane Library sugere que a conclusão seja estruturada em duas seções:

4.1.2.1 Implicações para prática clínica

O significado dos achados nas implicações para a prática clínica deve ser tão inequívoco quanto possível. Não devem ir além da evidência que foi revisada e ser justificada pelos dados apresentados na revisão. Por exemplo, Weller e cols. em sua revisão sistemática de estudos de coorte avaliando o uso de preservativos para redução da transmissão do vírus HIV em heterossexuais, concluíram que o uso consistente de preservativos reduz a transmissão de HIV e destacaram que “uso consistente” foi definido como uso de preservativo masculino em todas as relações sexuais via vaginal⁸⁰.

4.1.2.2 Implicações para pesquisas futuras

Se a metanálise foi julgada como conclusiva, isto deve ser relatado na conclusão. Metanálises que não demonstram haver diferenças entre os grupos, não necessariamente significam que não existam diferenças entre eles. A depender da imprecisão da medida de associação (evidenciada pelo intervalo de confiança), a metanálise pode não ter demonstrado diferença por não ter atingido um tamanho de amostra ou número de eventos suficientes que lhe confira poder estatístico adequado para evidenciar tais diferenças. Nesses casos, é importante explorar essa alternativa, enfatizando a necessidade de novos estudos para esclarecer a questão.

4.2 Utilização da diretriz de redação do manuscrito – MOOSE

Um grupo de 27 especialistas propôs um *checklist* metodológico para a redação de manuscritos reportando revisões sistemáticas e meta-análises de estudos observacionais denominado *Meta-analysis of observational studies in epidemiology* (MOOSE)².

O *checklist* MOOSE contém especificações para redação da introdução, estratégia de busca, métodos, resultados, discussão e conclusão. Este *checklist* objetiva a melhora da utilidade e do relato de revisões sistemáticas e meta-análises de estudos observacionais para autores, revisores, editores, leitores e tomadores de decisão clínica.

A revisão sistemática de estudos observacionais é aparentemente similar na metodologia aplicada às revisões sistemáticas de ensaios clínicos randomizados, mas difere por natureza, principalmente por ser menos segura e definitiva em relação à medida de associação sumária. No entanto, pode fornecer uma ferramenta útil na compreensão e quantificação das fontes de variabilidade entre os resultados dos estudos disponíveis na literatura. O Anexo E apresenta a tradução simples (não validada) seguida do instrumento original.

5.1 Avaliação da qualidade da evidência – sistema GRADE

O sistema Grading the quality of evidence and the strength of recommendations (GRADE) foi elaborado com o objetivo de estabelecer critérios claros para avaliar a qualidade de evidências científicas e formular recomendações com base nos achados.

O sistema GRADE inclui avaliação da qualidade da evidência, com posterior definição da direção e força das recomendações. Esta diferenciação entre a qualidade e a força da recomendação é uma das vantagens do sistema GRADE em relação aos outros sistemas de graduação da qualidade das evidências. Nem todos os sistemas de classificação diferenciam decisões de acordo com a qualidade da evidência ou a força de recomendações. Evidência de alta qualidade não implica necessariamente em recomendações fortes, e fortes recomendações podem surgir a partir de evidências de qualidade não ideal.

O sistema GRADE gradua a qualidade da evidência em quatro níveis: alta, moderada, baixa ou muito baixa, conforme Quadro 9. O GRADE tem sido utilizado amplamente por conceituadas organizações como a Organização Mundial da Saúde, American College of Physicians, UpToDate e Cochrane Collaboration.

Quadro 9 – Sistema GRADE para avaliação da qualidade da evidência⁸¹

Qualidade	Definição	Metodologia
Alta	Estamos muito confiantes de que o verdadeiro efeito está próximo da medida de efeito encontrada.	ECRs bem delineados e executados, produzindo resultados consistentes e aplicáveis e/ou estudos observacionais muito bem delineados e com grandes estimativas de efeito.
Moderada	Estamos moderadamente confiantes na medida de associação: é provável que o verdadeiro efeito esteja próximo da medida de efeito encontrada, mas há possibilidade de que seja substancialmente diferente.	ECRs com grandes limitações ou estudos observacionais bem delineados com grandes estimativas de efeito.
Baixa	Nossa confiança na estimativa de efeito é limitada: o verdadeiro efeito pode ser substancialmente diferente da medida de efeito.	Estudos observacionais bem delineados e, ocasionalmente, ECRs* com limitações importantes.
Muito Baixa	Temos muito pouca confiança na estimativa de efeito: é provável que o verdadeiro efeito seja substancialmente diferente da medida de efeito encontrada.	Estudos observacionais mal controlados e observações clínicas não sistemáticas (séries ou relatos de casos).

Fonte: (Balsheem e cols., 2011, adaptado)⁸¹.

*ECR = ensaio clínico randomizado

5.1.1 O sistema GRADE e a força da recomendação

A força da recomendação baseia-se no grau de confiança no equilíbrio entre efeitos desejáveis e indesejáveis de uma exposição. Se os efeitos desejáveis claramente excedem os efeitos indesejáveis ou vice-versa, a recomendação é considerada forte. De outro lado, se os efeitos desejáveis e indesejáveis estão em estreito equilíbrio, define-se uma fraca recomendação.

Além do equilíbrio entre efeitos desejáveis e indesejáveis, outros fatores que podem influenciar a força da recomendação são: qualidade da evidência, importância relativa e risco basal dos desfechos, magnitude da medida de associação e custos.

5.1.2 Como graduar a qualidade da evidência

A qualidade da evidência refere-se ao grau de confiança em uma determinada medida de associação. De acordo com o GRADE, a qualidade da evidência pode ser classificada como alta, moderada, baixa ou muito baixa qualidade.

Ensaio clínico randomizado são inicialmente categorizados como evidência de “alta qualidade”, mas a confiança na evidência pode ser reduzida por vários motivos, incluindo: limitações metodológicas do estudo, inconsistência dos resultados, evidências indiretas, imprecisão, suspeita de viés de publicação (Quadro 10). Por outro lado, quando o sistema GRADE é utilizado para graduar as evidências provenientes de estudos observacionais, a qualidade da evidência se inicia como de “baixa qualidade”. Uma elevação nos níveis da graduação é permitida se há grande magnitude do efeito, se existe evidência de uma relação dose-resposta, ou se todos os vieses plausíveis dos estudos observacionais contribuem para subestimar um aparente efeito de tratamento (Quadro 11).

Quadro 10 – Fatores que diminuem a qualidade da evidência⁸²

Qualidade	Definição
Limitação Metodológica	Viés de seleção, viés de aferição diferencial, perdas de seguimento.
Inconsistência	Heterogeneidade entre os estudos, evidenciada principalmente pelo teste de inconsistência de Higgins.
Evidência Indireta	A questão abordada não é respondida diretamente pelos estudos disponíveis devido a diferenças na população, nas exposições, comparadores ou desfechos.
Imprecisão	Intervalos de confiança largos (imprecisão quanto ao verdadeiro efeito da intervenção).
Viés de Publicação	Tendência de publicação de estudos apenas com resultados positivos, em idioma inglês, em revistas indexadas no MEDLINE. Disponibilidade exclusivamente de estudos pequenos é indício de maior risco de viés de publicação.

Fonte: Adaptado de Guyatt e cols.⁸²

Quadro 11 – Fatores que aumentam a qualidade da evidência⁷¹

Qualidade	Definição
Grande magnitude do efeito	Quando a magnitude do efeito é muito grande, o efeito observado dificilmente poderá ser atribuído aos potenciais confundidores.
Potenciais variáveis de confusão subestimam o efeito	Há situações em que as variáveis de confusão e outros vieses estão relacionados unicamente em reduzir a estimativa do efeito.
Gradiente dose-resposta	A presença do gradiente dose-resposta aumenta nossa confiança nas estimativas de estudos observacionais.

Fonte: (Guyatt, 2008)⁸².

1. BRASIL. Ministério da saúde. **Diretrizes Metodológicas. Elaboração de revisão sistemática e metanálise de ensaios clínicos randomizados**. Brasília: Editora do Ministério da Saúde, 2012.
2. STROUP, D. F. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE). **The Journal of the American Medical Association**, Chicago, v. 283, n. 15, p. 2008-2012, 2000.
3. VON, E. E. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. **British Medical Journal**, London, v. 355, n. 7624, p. 806-808, 2007.
4. GUYATT, G.; EVIDENCE BASED MEDICINE WORKING GROUP. **Users' guides to the medical literature: a manual for evidence-based clinical practice**. Chicago: American Medical Association, 2002.
5. COOK, D. J.; MULROW, C. D.; HAYNES, R. B. Systematic reviews: synthesis of best evidence for clinical decisions. **Annals of Internal Medicine**, Philadelphia, v. 126, n. 5, p. 376-380, 1997.
6. STROUP, D. F. et al. Characteristics of meta-analyses submitted to a medical journal. In: **Congress of Biomedical and Global Comm**. 1997.
7. WALKER, E.; HERNANDEZ, A. V.; KATTAN, M. W. Meta-analysis: Its strengths and limitations. **Cleveland Clinic Journal of Medicine**, Lyndhurst, v.75, n. 6, p. 431-439, 2008.
8. FLETCHER, R.; FLETCHER, S. **Clinical Epidemiology: The Essentials**. 4. ed. Philadelphia: Lippincott Williams & Wilkins, 2005.
9. HULLEY, S. et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. **The Journal of the American Medical Association**, Chicago, v.280, n. 7, p. 605-613, 1998.
10. YANG, D. et al. Effect of hormone replacement therapy on cardiovascular outcomes: a meta-analysis of randomized controlled trials. **PLoS One**, San Francisco, v. 8, n. 5, p. e62329, 2013.
11. AZIZ, F. Coronary Artery Disease in Women: An Unsolved Dilemma. **Journal of Clinical Medicine Research**, Quebec, v. 6, n. 2, p. 86-90, 2014.

12. GROENWOLD, R. H. et al. Poor quality of reporting confounding bias in observational intervention studies: a systematic review. **Annals of Epidemiology**, Philadelphia, v. 18, n. 10, p. 746-751, 2008.
13. MALTA, M. Iniciativa STROBE: subsídios para a comunicação de estudos observacionais. **Revista de Saúde Pública**, São Paulo, v. 44, n. 3, p. 559-565, 2010.
14. RICHARDSON, W. S. et al. The well-built clinical question: a key to evidence-based decisions. **ACP Journal Club**, Philadelphia, v. 123, n. 3, p. A12-A13, 1995.
15. FINEOUT-OVERHOLT, E. J. L. TEACHING EBP: asking searchable, answerable clinical questions. **Worldviews on Evidence-Based Nursing**, Malden, v. 2, n. 3, p. 157-160, 2005.
16. DEVEREAUX, P. J. Association between postoperative troponin levels and 30-day mortality among patients undergoing noncardiac surgery. **The Journal of the American Medical Association**, Chicago, v. 307, n. 21, p. 2295-2304, 2012.
17. THE COCHRANE COLLABORATION. **Cochrane Library**. Disponível em: <<http://www.thecochranelibrary.com/view/0/index.html>>. Acesso em: 10 out. 2013.
18. BIBLIOTECA VIRTUAL EM SAÚDE. BVS. **Biblioteca Cochrane**. Disponível em: <<http://www.bireme.br/php/index.php>>. Acesso em: 10 out. 2013.
19. UNIVERSITY OF YORK. **Centre for Reviews and Dissemination**. CRD . Disponível em: <<http://www.york.ac.uk/inst/crd/>>. Acesso em: 16 out. 2013.
20. PUBMED. **Clinical Queries**. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/clinical>>. Acesso em: 22 out. 2013.
21. REDE BRASILEIRA DE AVALIAÇÃO DE TECNOLOGIAS EM SAÚDE. **SISREBRATS**. Disponível em: <<http://189.28.128.101/rebrats/visao/sociedade/estudo.cfm>>. Acesso em: 23 out. 2013.
22. HIGGINS, J. P. T.; GREEN, S. (Ed.). **Cochrane Handbook for Systematic Reviews of Interventions** [internet]. Version 5.1.0. Australia (AUS): Department of Health and Ageing, Austrália [revisado em 2011; citado em 2013 Out 10]. The Cochrane Collaboration. Disponível em: <www.cochrane-handbook.org>. Acesso em: 23 out. 2013.
23. MEDICAL Journal editors take hard line on drug research. **The Washington Post**, Washington DC, 10 set. 2004. Disponível em: <<http://www.smh.com.au/articles/2004/09/09/1094530773888.html>>. Acesso em: 10 out. 2013.

24. UNIVERSITY OF YORK. **International Prospective Register of Ongoing Systematic Reviews** (Prospero). Disponível em: <<http://www.crd.york.ac.uk/PROSPERO/internet>>. Acesso em: 15 out. 2013.
25. BIBLIOTECA VIRTUAL EM SAÚDE. BVS. **Lilacs**. Disponível em: <<http://lilacs.bvsalud.org/>>. Acesso em: 10 out. 2013.
26. GOLDER, S. Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. **Health Information and Libraries Journal**, Malden, v. 23, n. 1, p. 3-12, 2006.
27. THOMSON REUTERS SCIENTIFIC. **Web of Science**. Disponível em: <<http://thomsonreuters.com/thomson-reuters-web-of-science/>>. Acesso em: 15 out. 2013.
28. EBSCO. Cinhal. **Cumulative Index to Nursing and Allied Health Literature**. Disponível em: <<http://www.ebscohost.com/biomedical-libraries/the-cinahl-database>>. Acesso em: 17 out. 2013.
29. WILKINS, T.; GILLIES, R. A.; DAVIS, K. EMBASE versus MEDLINE for family medicine searches: can MEDLINE searches find the forest or tree?. **Canadian Family Physician**, Ontario, v. 51, p. 848-849, 2005.
30. BIBLIOTECA VIRTUAL EM SAÚDE. **Scientific Electronic Library Online** (SciELO). Disponível em: <<http://www.scielo.org/index.php?lang=pt>>. Acesso em: 20 set. 2013.
31. CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Disponível em: <<http://www.capes.gov.br/>>. Acesso em: 21 out. 2013.
32. PROQUEST. **Dissertations & Theses Database**. Disponível em: <<http://www.proquest.com/en-US/catalogs/databases/detail/pqdt.shtml>>. Acesso em: 26 out. 2013.
33. TRIP DATABASE. **Trip Database – clinical search engine**. Disponível em: <<http://www.tripdatabase.com/>>. Acesso em: 20 set. 2013.
34. NATIONAL INSTITUTES OF HEALTH. **Clinical Trials**. Disponível em: <<http://clinicaltrials.gov/>>. Acesso em: 20 set. 2013.
35. ISRCTN. **International Standard Randomised Controlled Trial Number Register**. Disponível em: <<http://www.isrctn.org/>>. Acesso em: 10 out. 2013.
36. BRASIL. Ministério da Saúde. **Registro Brasileiro de Ensaios Clínicos** (ReBEC). Disponível em: <<http://www.ensaiosclinicos.gov.br/>>. Acesso em: 20 set. 2013.

37. ANNALS OF SURGICAL INNOVATION AND RESEARCH. **BMC Central Proceedings**. Disponível em: <<http://www.biomedcentral.com/bmcproc>>. Acesso em: 10 out. 2013. p. 23.
38. BIOMED CENTRAL'S MEETING. **BMC Meeting Abstracts**. Disponível em: <<http://www.biomedcentral.com/bmcmeetingabstracts/>>. Acesso em: 10 out. 2013.
39. PUBMED. **Clinical Queries**. Disponível em: <http://www.ncbi.nlm.nih.gov/books/NBK3827/#pubmedhelp.Clinical_Queries_Filters>. Acesso em: 10 out. 2013.
40. HAYNES, R. B. et al. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. **Journal of the American Medical Informatics Association**, London, v. 1, n. 6, p. 447-458, 1994.
41. HIGGINS, J. P. T.; GREEN, S. (Ed.). **Cochrane Handbook for Systematic Reviews of Interventions**. New York: Wiley-Blackwell, Cap. 10, p. 295-333. 2011.
42. COHEN, J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. **Psychological bulletin**, Washington DC, v. 70, n. 4, p. 213-220, 1998.
43. LEE, C. H. et al. Carcinogenetic impact of alcohol intake on squamous cell carcinoma risk of the oesophagus in relation to tobacco smoking. **European Journal of Cancer**, Philadelphia, v. 43, n. 7, p. 1188-1199, 2007.
44. WEST, S. et al. **Systems to rate the strength of scientific evidence**. Evidence report technology assessment (Summary), v. 47, p. 1-11, 2002.
45. DEEKS, J. J. et al. Evaluating non-randomised intervention studies. **Health Technology Assessment**, Havant, v. 7, n. 27, p.iii-173, 2003.
46. SANDERSON, S.; TATT, I. D.; HIGGINS, J. P. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. **International Journal of Epidemiology**, Oxford, v. 36, n. 3, p. 666-676, 2007.
47. DOWNS, S.H.; BLACK, N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. **Journal of Epidemiology and Community Health**, London, v. 52, n. 6, p. 377-384, 1998.
48. VON, E.; ROLLIN, A. Ottawa Hospital Research Institute. **The Newcastle-Ottawa Quality Assessment Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses**. Disponível em: <http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp 2013>. Acesso em: 10 out. 2013.

49. ARENZ, S. Breast-feeding and childhood obesity-a systematic review. **International Journal of Obesity**, London, v. 28, n. 10, p. 1247-1256, 2004.
50. HORTA, B. L. Maternal smoking and the risk of early weaning: a meta-analysis. **American Journal of Public Health**, Birmingham, v. 91, n. 2, p. 304-307, 2001.
51. HARDER, T. Duration of breastfeeding and risk of overweight: a meta-analysis. **American Journal of Epidemiology**, Oxford, v. 162, n. 5, p. 397-403, 2005.
52. Di, P. C. Four-fold table cell frequencies imputation in meta analysis. **Statistics in Medicine**, Malden, v. 25, n. 13, p. 2299- 2322, 2006.
53. VERONIKI, A. A. Reconstructing 2 x 2 contingency tables from odds ratios using the Di Pietrantonj method: difficulties, constraints and impact in meta-analysis results. **Research Synthesis Methods**, Malden, v. 4, n. 1, p. 78-94, 2012.
54. INSTITUTE FOR STATISTICS AND MATHEMATICS. **Programa Estatístico “R”**. Disponível em: <<http://www.r-project.org/>>. Acesso em: 10 out. 2013.
55. COMPREHENSIVE META-ANALYSIS. **Meta-Analysis Version 5.3**. Disponível em: <<http://www.meta-analysis.com/index.php?gclid=COSSpL2U1rwCFdTm7AodsVkArA>> Acesso em: 10 out. . 2013.
56. META-ANALYSIS MADE EASY. **Metaanalysis with Interactive Explanations MIX 2.0**. Disponível em: <<http://www.meta-analysis-made-easy.com/>>. Acesso em: 10 out. 2013.
57. COMPREHENSIVE META-ANALYSIS. **Meta-Analysis**. Disponível em: <<http://www.meta-analysis.com/index.php>>. Acesso em: 10 out. 2013.
58. ROSENBERG, M. S. **MetaWin**. Disponível em: <<http://www.metawinsoft.com/>>. Acesso em: 20 set. 2013.
59. SOFT32. **Weasyma**. Disponível em:<<http://weasy-ma.soft32.com/>>. Acesso em: 20 set. 2013.
60. SAS. **Statistical Analysis System**. Disponível em: <<http://www.sas.com/offices/latinamerica/brazil/>>. Acesso em: 10 out. 2013.
61. STATA. **Data analysis and statistical software**. Disponível em: <<http://www.stata.com/>>. Acesso em: 20 set. 2013.
62. WINBUGS. **Bayesian inference using gibbs smpling**. Disponível em: <<http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>>. Acesso em: 20 set. 2013.

63. The COCRHANE COLLABORATION. **Review Manager**. Disponível em: <<http://www.cochrane.org/editorial-and-publishing-policy-resource/review-manager-revman>>. Acesso em: 20 set. 2013.
64. DICKERSIN, K. et al. The existence of publication bias and risk factors for its occurrence. **The Journal of the American Medical Association**, Chicago, v. 263, n. 10, p. 1385-1389, 1990.
65. DICKERSIN, K. et al. Publication bias and clinical trials. **Controlled Clinical Trials**, Maryland Heights, v. 8, n. 4, p. 343-353, 1987.
66. STERNE, J. A. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. **Journal of Clinical Epidemiology**, Maryland Heights, v. 54, n. 10, p. 1046-1055, 2001.
67. EGGER, M. et al. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. **Health Technology Assessment**, Winchester, v. 7, n. 1, p. 1-76, 2003.
68. SCHULZ, K. F. et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. **The Journal of the American Medical Association**, Chicago, v. 273, n. 5, p. 408-412, 1995.
69. TERRIN, N. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. **Journal of Clinical Epidemiology**, Maryland Heights, v. 58, n. 9, p. 894-901, 2005.
70. EGGER, M. Bias in meta-analysis detected by a simple, graphical test. **British Medical Journal**, London, v. 315, n. 7109, p. 629-634, 1997.
71. PETERS, J. L. Comparison of two methods to detect publication bias in meta-analysis. **The Journal of the American Medical Association**, Chicago, v. 295, n. 6, p. 676-680, 2006.
72. DUVAL, S. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. **Biometrics**, Arlington, v. 56, n. 2, p. 455-463, 2000.
73. HARDY, R.J. Detecting and describing heterogeneity in meta-analysis. **Statistics in Medicine**, Malden, v. 17, n. 8, p. 841-856, 1998.
74. THOMPSON, S. G. Why sources of heterogeneity in meta-analysis should be investigated. **British Medical Journal**, London, v. 309, n. 6965, p.1351-1355, 1994.

75. HIGGINS, J. P. Quantifying heterogeneity in a meta-analysis. **Statistics in Medicine**, Malden, v. 21, n. 11, p. 1539-1558, 2002.
76. HORTA, B. L. Maternal smoking and the risk of early weaning: a meta-analysis. **American Journal of Public Health**, Birmingham, 2001.
77. ARENZ, S. Breast-feeding and childhood obesity--a systematic review. **International Journal of Obesity and Related Metabolic Disorders**, London, v. 28, n. 10, p. 1247-1256, 2004.
78. KARTHIKEYAN, G. Is a pre-operative brain natriuretic peptide or N-terminal pro-B-type natriuretic peptide measurement an independent predictor of adverse cardiovascular outcomes within 30 days of noncardiac surgery? a systematic review and meta-analysis of observational studies. **Journal of the American College of Cardiology**, Amsterdam, v. 54, n. 17, p.1599-1606, 2009.
79. DOCHERTY, M. The case for structuring the discussion of scientific papers. **British Medical Journal**, London, v. 318, n. 7193, p. 1224-1225,1999.
80. WELLER, S. C.; DAVIS-BEATY, K. **Condom effectiveness in reducing heterosexual HIV transmission**. Cochrane Database of Systematic Reviews. The Cochrane Collaboration. Disponível em: <<http://apps.who.int/whl/reviews/CD003255.pdf>>. Acesso em: 10 out. 2013.
81. BALSHEM, H. et al. GRADE guidelines: 3. Rating the quality of evidence. **Journal of Clinical Epidemiology**, Maryland Heights, v. 64, n. 4, p. 401-406, 2011.
82. GUYATT, G.H. Rating quality of evidence and strength of recommendations GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. **British Medical Journal**, London, v. 336, n. 7650, p. 924-926, 2008.
83. BROUGHTON, T.; SINGTON, J.; BEALES, I. L. Statin use is associated with a reduced incidence of colorectal cancer: a colonoscopy-controlled case-control study. **BMC Gastroenterology**, London, v. 24, n.12, p. 36, 2012.
84. LYTRAS, T.; NIKOLOPOULOS, G.; BONOVAS, S. Statins and the risk of colorectal cancer: An updated systematic review and meta-analysis of 40 studies. **World Journal of Gastroenterology**, Hong Kong, v. 20, n. 7, p. 1858-1870, 2014.
85. RAUTIAINEN, S. et al. Vitamin C supplements and the risk of age-related cataract: a population-based prospective cohort study in women. **American Journal of Clinical Nutrition**, Bethesda, v. 91, n. 2, p. 487-493, 2010.

86. ZHAO, L. Q.; LI, L. M.; ZHU, H. The effect of multivitamin/mineral supplements on age-related cataracts: a systematic review and meta-analysis. **Nutrients**, Basel, v. 6, n. 3, p. 931-949, 2014.
87. SCHULZ, K. F. et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. **The Journal of the American Medical Association**, Chicago, v. 273, n. 5, p. 408-412, 1995.
88. PARMAR, M. K.; TORRI, V.; STEWART, L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. **Statistic in Medicine**, Malden, v. 17, n. 24, p. 2815-2834, 1998.
89. KARTHIKEYAN, G. Is a pre-operative brain natriuretic peptide or N-terminal pro-B-type natriuretic peptide measurement an independent predictor of adverse cardiovascular outcomes within 30 days of noncardiac surgery? a systematic review and meta-analysis of observational studies. **Journal of the American College of Cardiology**, San Diego, v. 54, n. 17, p.1599-1606, 2009.
90. PLOS MEDICINE. Disponível em: < <http://www.plosmedicine.org/> >. Acesso em: 10 out. 2013.
91. ANNALS OF INTERNAL MEDICINE. **Epidemiology**, Disponível em: <<http://annals.org/> internet>. Acesso em: 10 out. 2013.
92. WILCOX, A. J. **Epidemiology**, Disponível em: <<http://journals.lww.com/epidem/pages/default.aspx>>. Acesso em: 10 out. 2013.
93. UNIVERSITAT BERN. **STrengthening the Reporting of OBservational studies in Epidemiology** (Strobe). Disponível em: <<http://www.strobe-statement.org/>>. Acesso em: 20 set. 2013.

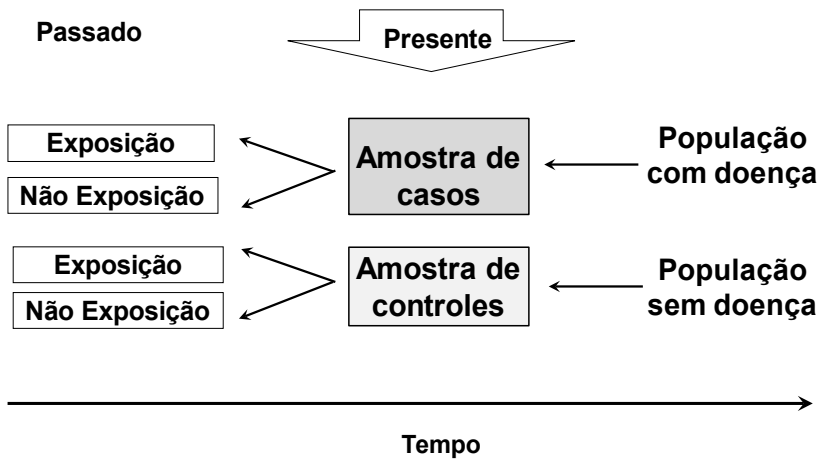


APÊNDICES

Apêndice A – Principais tipos de estudos observacionais para avaliação de risco e prognóstico

Estudos caso-control			
Definição	Etapas	Vantagens	Desvantagens
É um estudo observacional no qual os indivíduos são selecionados de acordo com a presença (casos) ou não (controles) da doença. Da comparação dos grupos, procura-se definir fatores e atributos que os diferenciem (possivelmente relacionados com a “causa” da doença).	- Seleção de casos incidentes; - Seleção de controles na mesma base populacional que os casos; - Pareamento; - Informação sobre a exposição; - A medida de associação entre a “causa” (exposição) e o “efeito” (doença) é dada pelo <i>odds ratio</i>.	- Geralmente de baixo custo e menor tempo quando comparado com os estudos de coorte; - Pode ser ideal para estudo de doenças raras ou com longo período de latência; - Permitem o estudo de múltiplas exposições.	- A escolha e o recrutamento da população-controle é um desafio (viés de seleção/viés de memória); - O risco é estimado de maneira indireta pelo <i>odds ratio</i>.

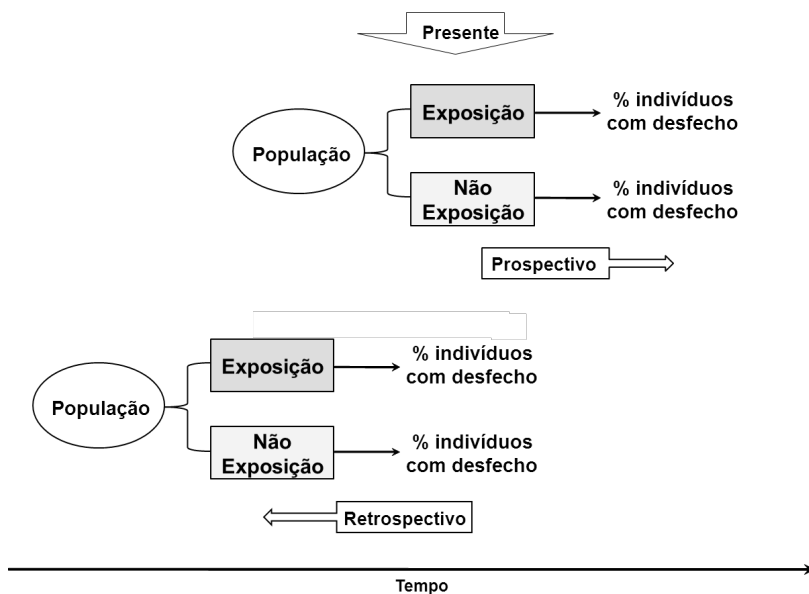
Estudo de Caso-Control



Exemplo: Broughton e cols. conduziram um estudo de caso-control avaliando o uso de estatinas em pacientes com câncer colorretal. Os 101 casos foram selecionados entre os pacientes sintomáticos que tiveram seus diagnósticos confirmados por meio de colonoscopia com análise histológica. Os 132 controles foram selecionados entre pacientes também sintomáticos, submetidos à colonoscopia sem diagnóstico de câncer colorretal. Os autores utilizaram questionários estruturados e o prontuário médico para coletar dados sobre o uso de estatinas. Entre os pacientes que utilizaram estatina previamente, 20 pertenciam ao grupo caso e 68 ao grupo controle. O *odds ratio* ajustado (para covariáveis idade, gênero, diabetes, uso de aspirina, metformina e consumo de álcool semanal) foi de 0,43 (intervalo de 95% de confiança de 0,25-0,80)⁸³. O estudo de Broughton foi incluído na revisão sistemática de Lytras e cols. avaliando a associação do uso de estatinas e o risco de câncer colorretal que incluiu 19 estudos de caso-control (além de outros desenhos de estudos). A metanálise dos estudos de caso-control demonstrou um *odds ratio* de 0,92 (intervalo de 95% de confiança de 0,87-0,98)⁸⁴.

Estudos de coorte			
Definição	Etapas	Vantagens	Desvantagens
Compara a experiência de um grupo exposto e outro não exposto ao longo do tempo, para a identificação dos efeitos da exposição na incidência do evento de interesse.	- Identificar grupos de indivíduos expostos e não expostos; - Observar cada coorte ao longo do tempo para a avaliação do desenvolvimento da doença nos grupos estudados; - Comparar os riscos de surgimento da doença entre os grupos de expostos e não-expostos.	- Relação temporal; - Cálculo direto da taxa de incidência; - Útil em exposição rara; - Minimiza viés de seleção; - Pode estudar múltiplas doenças; - Pode acessar a exposição em diferentes pontos no tempo.	- Logística e viabilidade; - Custo e tempo; - Dificuldade de manter retenção de pacientes em estudos de longo período de seguimento, sendo que a perda pode comprometer a validade dos resultados.

Estudo de COORTE



Exemplo: Em 2010, Rautiainen e cols. publicaram um estudo de coorte avaliando a relação entre suplementos polivitamínicos contendo vitamina C e a incidência de catarata em 24.593 mulheres com idade entre 49 e 83 anos. As participantes foram seguidas por 8,2 anos (184.698 pessoas-ano) e 2.497 novos casos de catarata foram diagnosticados. A análise multivariada demonstrou um *hazard ratio* de 1,09 (intervalo de 95% de confiança 0,94 – 1,25) para usuárias de suplementos contendo vitamina C em relação às não usuárias⁸⁵. Esse estudo foi posteriormente incluído na revisão sistemática e metanálise conduzida por Zhao e cols. que também não encontraram associação entre suplementos polivitamínicos e catarata⁸⁶.

Apêndice B – Exemplo de sensibilização de busca utilizando os “entry terms” do termo MeSH para “brain natriuretic peptide”

Ex.: Termo MeSH + *entry terms*

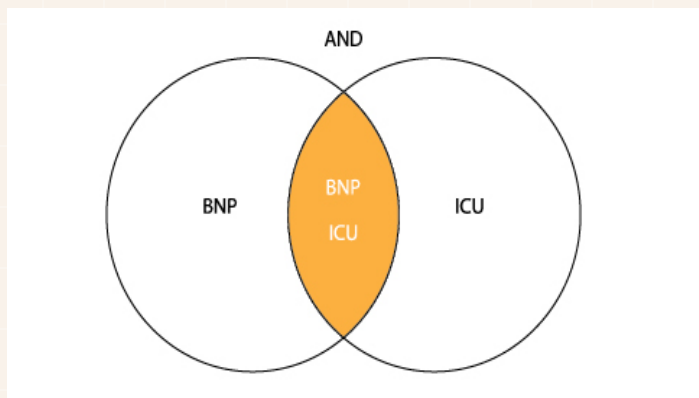
“Natriuretic Peptide, Brain”[Mesh] OR (Peptide, Brain Natriuretic) OR (Brain Natriuretic Peptide) OR (BNP-32) OR (BNP 32) OR (Nesiritide) OR (B-Type Natriuretic Peptide) OR (Natriuretic Peptide, B-Type) OR (BNP Gene Product) OR (Type-B Natriuretic Peptide) OR (Natriuretic Peptide, Type-B) OR (Type B Natriuretic Peptide) OR (Natriuretic Peptide Type-B) OR (Natriuretic Peptide Type B) OR (Natriuretic Factor-32) OR (Natriuretic Factor 32) OR (Brain Natriuretic Peptide-32) OR (Brain Natriuretic Peptide 32) OR (Natriuretic Peptide-32, Brain) OR (Peptide-32, Brain Natriuretic) OR (Ventricular Natriuretic Peptide, B-type) OR (Ventricular Natriuretic Peptide, B type) OR (Natrecor).

Apêndice C – Construção da estratégia de busca: símbolos e operadores booleanos

Os operadores booleanos “AND”, “OR” e “NOT” são responsáveis por fazer a ligação dos termos de busca e representam as seguintes ações:

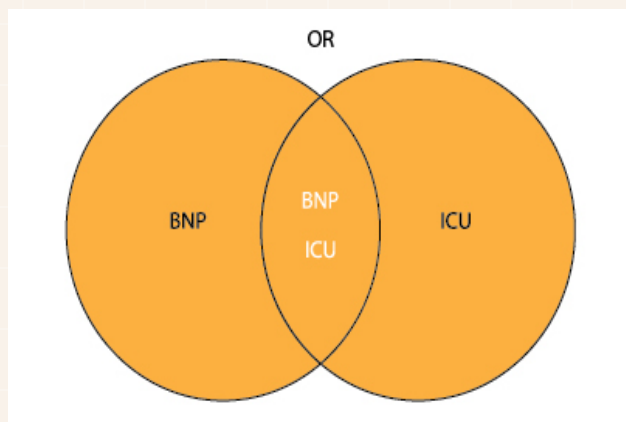
AND: representa a intersecção dos termos. Quando utilizado, informamos à base que queremos um artigo que contenha, necessariamente, as duas palavras. Por exemplo, se queremos encontrar artigos que tenham avaliado o peptídeo natriurético encefálico como marcador prognóstico de doença cardiovascular em pacientes de unidade de terapia intensiva, devemos utilizar o seguinte arranjo: “Natriuretic Peptide, Brain” AND “Intensive Care Units”. Assim, os artigos recuperados por esta busca conterão estudos que só tenham avaliado peptídeo natriurético encefálico em pacientes internados em unidade de terapia intensiva. A ordem dos termos não altera os resultados. A Figura C.1 abaixo ilustra o conceito de busca utilizando o operador booleano AND:

Figura C.1 – Operador booleano AND



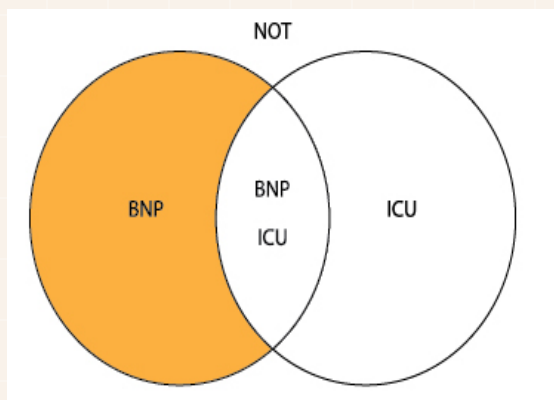
Fonte: elaboração própria.

OR: representa a soma, a união dos termos. Ele permite recuperar artigos que contenham tanto uma, como a outra expressão, inclusive artigos que contenham ambos os assuntos. No exemplo ilustrado anteriormente, ao utilizarmos o operador OR “Natriuretic Peptide, Brain” OR “Intensive Care Units”, recupera-se todos os artigos com o assunto Natriuretic Peptide, Brain, todos com o assunto “Intensive Care Units” e também com ambos. A ordem dos termos não altera os resultados, assim como no operador AND. A Figura C.2 abaixo ilustra o conceito de busca utilizando o operador booleano OR:

Figura C.2 – Operador booleano OR

Fonte: elaboração própria.

NOT: O NOT é o operador booleano que exclui o termo subsequente. Pelo exemplo anterior, o arranjo “Natriuretic Peptide, Brain” NOT “Intensive Care Units” recupera somente artigos com o assunto “Natriuretic Peptide, Brain”, ou seja, exclui todos os que incluem “Intensive Care Units”, inclusive os artigos que contêm ambos os termos. Neste caso, a ordem dos termos altera os resultados. A Figura C.3 abaixo ilustra o conceito de busca utilizando o operador booleano NOT:

Figura C.3 – Operador booleano NOT

Fonte: elaboração própria.

Aplicando estas definições para a questão de pesquisa abaixo, o Quadro C.1 representa a construção da estratégia de busca e realização da pesquisa por evidências para a questão de pesquisa a seguir:

Questão de pesquisa: *Is the Brain Natriuretic Peptide a prognostic marker in critically ill patients?*

Exemplo de construção de estratégia de busca no MEDLINE:

Quadro C.1 – Exemplo de construção de estratégia de busca no Medline/Pubmed

MEDLINE/PUBMED: Utilização do descritor de assunto (termo Mesh) mais a sensibilização com a utilização dos “entry terms”.

População/condição: Pacientes de unidade de terapia intensiva

“Intensive Care Units”[Mesh] OR (Care Unit, Intensive) OR (Care Units, Intensive) OR (Intensive Care Unit) OR (Unit, Intensive Care) OR (Units, Intensive Care) OR “Intensive Care”[Mesh] OR (Care, Intensive) OR (Surgical Intensive Care) OR (Care, Surgical Intensive) OR (Intensive Care, Surgical) OR Critical Care”[Mesh] OR (Care, Critical) OR (Critical illness”) OR (“critically patients”) OR (critically ill patients).

AND

Intervenção: “Natriuretic Peptide, Brain”[Mesh] OR (peptide, brain natriuretic) OR (brain natriuretic peptide) OR (BNP-32) OR (BNP 32) OR (nesiritide) OR (b-type natriuretic peptide) OR (natriuretic peptide, b-type) OR (xnp gene product) OR (type-b natriuretic peptide) OR (natriuretic peptide, type-b) OR (type b natriuretic peptide) OR (natriuretic peptide type-b) OR (natriuretic peptide type b) OR (natriuretic factor-32) OR (natriuretic factor 32) OR (brain natriuretic peptide-32) OR (brain natriuretic peptide 32) OR (natriuretic peptide-32, brain) OR (peptide-32, brain natriuretic) OR (ventricular natriuretic peptide, b-type) OR (ventricular natriuretic peptide, b type) OR (natrecor).

AND

Controle: não incluir termos para o comparador.

AND

Desfecho: heart failure OR (Cardiac Failure) OR (Myocardial Failure) OR (Heart Failure, Left-Sided) OR (Heart Failure, Left Sided) OR (Left-Sided Heart Failure) OR (Left Sided Heart Failure) OR (Heart Failure, Right-Sided) OR (Heart Failure, Right Sided) OR (Right-Sided Heart Failure) OR (Right Sided Heart Failure) OR (Congestive Heart Failure) OR (Heart Failure, Congestive) OR (Heart Decompensation) OR (Decompensation, Heart).

Exemplo na LILACS:

(Unidades de Terapia Intensiva OR Unidades de Cuidados Intensivos OR Intensive Care Units OR Centro de Terapia Intensiva OR Centros de Terapia Intensiva OR CTI OR Unidade de Terapia Intensiva OR UTI) (Peptídeo Natriurético Encefálico OR Péptido Natriurético Encefálico OR Natriuretic Peptide, Brain OR Peptídeo Encefálico Natriurético OR Peptídeo Cerebral Natriurético OR Peptídeo Natriurético Cerebral OR Nesiritida)

Elaborar uma única versão de uma estratégia de busca geralmente não é possível. O revisor pode se deparar com uma quantidade grande de artigos, o que pode inviabilizar a seleção de tamanho quantitativo. Neste caso, terá que lançar mão de algum termo que deixe a busca mais específica.

Assim, para otimizar tempo, sugere-se que a pesquisa seja realizada “linha por linha”, ou seja, que para cada termo explorado se realize uma pesquisa individual. Com isto, torna-se possível realizar diversas combinações utilizando os operadores booleanos. Os termos podem ser copiados em arquivo *Word* para que, sempre que necessário rodar uma nova estratégia, possam ser “colados” na caixa de pesquisa da respectiva base.

Algumas ferramentas adicionais para construção da busca merecem ser destacadas. O uso de símbolos como * (asterisco), \$(cifrão), ? (ponto de interrogação), vai depender de cada base utilizada. O símbolo de truncamento no Medline/Pubmed e Embase é o *(asterisco), conforme exemplificado por: random*. Ao utilizar esta construção, estamos buscando artigos que contenham as seguintes palavras: “random”, “randomised”, “randomized” e “randomly”.

No EMBASE, o uso do ? (ponto de interrogação) representa a recuperação de mais de um termo, como mostra o exemplo: wom?n busca artigos que utilizem as palavras “woman” e “women”.

Apêndice D – Gerenciadores de referências: o que são e como utilizá-los

Um gerenciador de referências apresenta inúmeras funções, inclusive, gerenciar as referências na condução de uma revisão sistemática. Permite adicionar os resultados de todas as bases eletrônicas pesquisadas em um único arquivo, bem como adicionar referências manualmente gerando uma biblioteca pesquisável. Ainda, permite organizar de forma prática todas as potenciais citações a serem incluídas na revisão sistemática, classificando-as por ordem alfabética de título, nome do autor ou número de identificação muito útil na remoção de referências duplicadas.

Várias são as opções de gerenciadores de referências disponíveis no mercado. A tabela D.1 fornece as principais opções comerciais ou gratuitas, bem como a disponibilidade para os sistemas operacionais. Nessa tabela, segue a relação dos gerenciadores de referências mais comuns.

Tabela D.1 – Gerenciadores de referências

Software	Acesso	Windows	Mac OS X	Linux
Aigaion	Gratuito	Sim	Sim	Sim
Bebop	Gratuito	Sim	Sim	Sim
BibDesk	Gratuito	Não	Sim	Não
Biblioscape	Comercial	Sim	Não	Não
Bibus	Gratuito	Sim	Em teste	Sim
Bookends	Comercial	Não	Sim	Não
Citavi	Comercial	Sim	Não	Não
Connotea	Gratuito	Sim	Sim	Sim
Docear	Gratuito	Sim	Sim	Sim
EndNote	Comercial	Sim	Sim	Não
EndNote Web	Gratuito	Sim	Sim	Não
JabRef	Gratuito	Sim	Sim	Sim
Jumper 2.0	Gratuito	Sim	Sim	Sim
KBibTeX	Gratuito	Em teste	Em teste	Sim
Mendeley	Gratuito versão básica	Sim	Sim	Sim
Papers	Comercial	Sim	Sim	Não
Pybliographer	Gratuito	Parcial	Parcial	Sim
Qiqqa	Gratuito	Sim	Não	Não
refbase	Gratuito	Sim	Sim	Sim
RefDB	Gratuito	Sim	Sim	Sim

Continua

Conclusão

Reference Manager	Comercial	Sim	Não	Não
Referencer	Gratuito	Não	Não	Sim
RefWorks	Comercial	Sim	Sim	N/A
Scholar's Aid	Gratuito versão básica	Sim	Não	Não
Sente	Comercial	Não	Sim	Não
Wikindx	Gratuito	Sim	Sim	Sim
WizFolio	Gratuito versão básica	Sim	Sim	Sim
Zotero	Gratuito versão básica	Sim	Sim	Sim

Cada gerenciador de referências tem sua peculiaridade que deve ser explorada na ocasião de sua utilização. Para os gerenciadores mais comuns, o EndNote® e o Reference Manager®, a maioria das etapas fundamentais são comuns e criar uma biblioteca única requer poucas etapas.

Independente do gerenciador de referências utilizado, deve-se obter o resultado da busca em arquivo texto (txt) para posterior inserção no gerenciador. Assim, utiliza-se o comando “exportar” das bases eletrônicas. Para o Medline, exportam-se os resultados da busca clicando em “send to” (localizado no canto superior direito da tela de resultados) e, após, seleciona-se a opção “file” e formato “MEDLINE”. Então, clica-se no botão “Create file” para criar o arquivo de texto. A figura D.1 ilustra o procedimento:

Figura D.1 – Tela do Pubmed

The screenshot shows the PubMed.gov search interface. At the top, there's a navigation bar with 'NCBI', 'Resources', and 'How To'. Below it, the 'PubMed.gov' logo is displayed along with 'U.S. National Library of Medicine' and 'National Institutes of Health'. A search bar contains the text 'Myocardial Infarction' with '[Mesh]' in brackets. To the right of the search bar are links for 'RSS', 'Save search', 'Limits', 'Advanced search', and 'Help'. Below the search bar, there are buttons for 'Search' and 'Clear'. On the left side, there's a link for 'Display Settings' and a summary of results: 'Summary, 20 per page, Sorted by Recently Added'. In the center, the results are listed, starting with 'Myocardial infarction due to percutaneous coronary intervention' by Prasad A, Hermann J. The first result is circled in red. On the right side, there's a 'Send to' dropdown menu with options: 'File', 'Clipboard', 'Collections', 'E-mail', 'Order', and 'My Bibliography'. Below this, there's a 'Format' dropdown set to 'MEDLINE' and a 'Sort by' dropdown set to 'Recently Added'. At the bottom right, there's a 'Create File' button.

Results: 1 to 20 of 129292

1. [Myocardial infarction due to percutaneous coronary intervention](#).
 Prasad A, Hermann J.
 N Engl J Med. 2011 Feb 3;364(5):453-64. Review. No abstract available.
 PMID: 21288097 [PubMed - indexed for MEDLINE]
[Related citations](#)

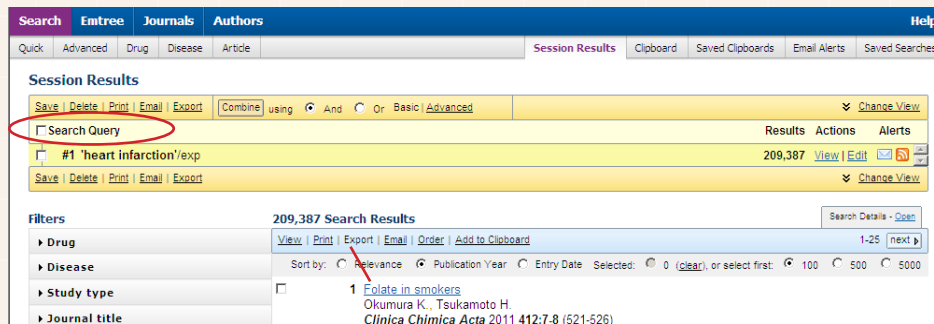
2. [Hospital readmission as an accountability measure](#).
 Axon RN, Williams MV.
 JAMA. 2011 Feb 2;305(5):504-5. No abstract available.
 PMID: 21285430 [PubMed - indexed for MEDLINE]

Format: MEDLINE
 Sort by: Recently Added
 Create File

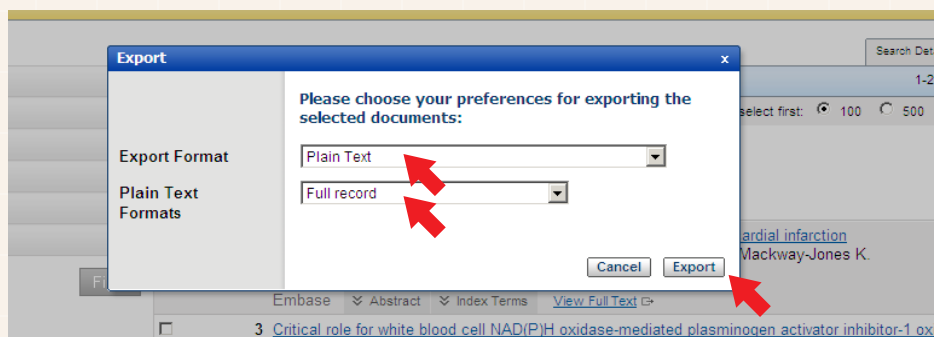
Fonte: PUBMED, 2013.

Para exportar os resultados da busca no EMBASE, clicar em “export” localizado na barra azul superior da tela de resultados. Então, escolhe-se o formato de exportação “Plain text” “Full record” e clicar em “export”, conforme Figuras D.2 e D.3.

Figura D.2 – Tela do EMBASE



Fonte: EMBASE, 2013.



Fonte: EMBASE, 2013.

O resultado da busca em cada base de dados deve ser exportado para um arquivo de texto para permitir a avaliação da elegibilidade. Utilizando um programa de gerenciador de referências, estes arquivos de texto podem ser somados por meio do comando “importar” do programa. Para isto, seleciona-se o filtro de importação que é específico para cada base de dados. Os gerenciadores possuem inúmeros filtros e dificilmente alguma base pesquisada não encontrará seu filtro de importação correspondente.

FORMULÁRIO DE ELEGIBILIDADE

Revisão sistemática e metanálise para avaliação do Peptídeo Natriurético Cerebral (BNP) como preditor prognóstico em pacientes internados em unidade de terapia intensiva

NOME DO REVISOR _____

IDENTIFICAÇÃO DO ARTIGO

Sobrenome do autor (se o autor pertencer a um grupo de pesquisa, indicar o nome do grupo) _____

Nome do Jornal (abreviação comumente utilizada, nome no PubMed ou nome completo do Jornal) _____

Ano de Publicação _____ Volume _____

Número da 1ª página _____

Número do artigo na base do "RefMan" _____

CRITÉRIOS DE ELEGIBILIDADE

Tipo de estudo

Estudo de coorte (prospectivo ou retrospectivo)?

() Sim () Não () Não está claro



Excluir

Participantes

Pacientes críticos internados em unidade de terapia intensiva? (Exceto unidade de cuidados cardiológicos/coronários exclusivos)

() Sim () Não () Não está claro



Excluir

Todos os participantes têm diagnóstico de doença arterial coronariana, embolia pulmonar, cirurgia cardíaca, insuficiência cardíaca ou choque cardiogênico?

() Sim () Não () Não está claro



Excluir

Exposição

Foi realizada dosagem de BNP na admissão dos pacientes na unidade de terapia intensiva?

() Sim () Não () Não está claro



Excluir

Desfecho

Os autores apresentaram medida de associação independente (análise multivariada) entre BNP e mortalidade?

() Sim () Não () Não está claro



Excluir

DECISÃO FINAL

() Incluir () Excluir () Não está claro → Entrar em contato com autores

APÊNDICE F – Cálculo e interpretação das medidas de associação

A) Medidas de associação para desfechos de dados dicotômicos

Para compreensão das definições das medidas de associação de desfechos dicotômicos, estes dados podem ser apresentados segundo uma tabela 2 x 2:

EXPOSIÇÃO	DESECHO			TOTAL
		Sim	Não	
	Sim	a	b	a + b
	Não	c	d	c + d
		a + c	b + d	

1) Risco absoluto (RA): É a probabilidade de desenvolver o desfecho em cada grupo. É matematicamente representado por $a/(a+b)$ e $c/(c+d)$.

2) Redução absoluta de risco (RAR): é a diferença de risco absoluto entre o grupo exposto e o grupo não exposto. É matematicamente representado por:

$$RAR = c/(c+d) - a/(a+b).$$

3) Risco relativo (RR): é o risco de eventos em indivíduos expostos, em relação ao risco em indivíduos não expostos, ou seja, $RR = [a/(a+b)] / [c/(c+d)]$.

4) Redução de risco relativo (RRR): É uma estimativa da proporção do risco basal que é removida ou aumentada pela exposição. Pode ser calculado de duas maneiras:

$$RRR = 1 - RR \text{ ou } RRR = RRA/RA \text{ (no grupo controle).}$$

5) Odds ratio (OR): $(a/b) / (c/d) = ad / bc$. O *odds ratio* é uma razão de chances. Por sua vez, chance é a probabilidade de um evento ocorrer sobre a probabilidade do evento não ocorrer.

A seguir, um exemplo numérico hipotético, para o cálculo dos riscos e chances:

	IAM	Não IAM	Total
Tabagista	80	20	100
Não Tabagista	20	80	100
Total	100	100	200

Exemplo hipotético para cálculo das estimativas de efeito

Risco e Risco Relativo:

Risco Absoluto dos expostos: $80/100 = 0,8$

Risco Absoluto dos não expostos: $20/100 = 0,2$

Razão de risco (risco relativo) = $0,8/0,2 = 4$

Interpretação: o risco de IAM dos expostos ao tabagismo é 4x o risco dos não expostos

Chance e Odds Ratio:

Chance dos expostos: $80/20 = 4$

Chance dos não expostos: $20/80 = 0,25$

Odds Ratio = $4/0,25 = 16$

Interpretação: A chance de ser tabagista e ter um evento de infarto é 16x a de quem não é tabagista ter um infarto.

Em relação às medidas de efeito de desfechos dicotômicos, alguns pontos importantes merecem ser destacados. Risco Relativo é, geralmente, mais intuitivamente compreensível para a maioria das pessoas⁸⁷. Apesar do “risco” ser comumente utilizado como sinônimo de “odds”, em *estatística* eles têm significados distintos e são calculados de formas diferentes. Quando essas diferenças são ignoradas, os resultados das metanálises podem ser mal interpretados. Odds pode ser convertido em risco, e o risco em odds, utilizando as fórmulas:

$$\text{Risco} = \frac{\text{odds}}{1 + \text{odds}} \quad \text{odds} = \frac{\text{risco}}{1 - \text{risco}}$$

A interpretação de odds é mais complexa que a de risco e para assegurar sua correta interpretação. Deve-se, primeiro, transformar odds em risco. De forma prática, a diferença entre odds e risco é pequena quando o evento é raro. Por exemplo, se o odds é 0,1, significa que uma pessoa terá o evento para cada 10 que não o terão, e, utilizando a fórmula, o risco será $0,1 / (1 + 0,1) = 0,091$.

Quando os eventos são comuns, as diferenças entre odds e risco passam a ser grandes. Por exemplo, o risco de 0,5 é equivalente a um odds de 1 e um risco de 0,95 é equivalente a um odds de 19.

Para os desfechos dicotômicos, as medidas de associação nos estudos devem ser apresentadas como razões, já que se tratam de comparações entre dois grupos. Assim como a chance (odds) difere de risco quando os eventos são comuns, o odds ratio também difere da razão de risco (ou risco relativo). Para intervenções que aumentam a chance de evento, o odds ratio é maior que o risco relativo, portanto, se tratados como sinônimos, a interpretação do resultado tenderá a superestimar o

efeito da intervenção. Para intervenções que reduzem a chance de evento, o *odds ratio* tende a ser menor que o risco relativo, mais uma vez superestimando o efeito da intervenção.

6) *Hazard ratio*:

Hazard ratio é uma medida de efeito utilizada para análises de sobrevivência ou análises do tipo “tempo até o evento”. Este tipo de estimativa não leva em conta apenas se o desfecho ocorreu ou não, mas também o ponto no tempo em que o evento ocorreu, ao longo do seguimento do estudo.

O cálculo da metanálise para *hazard ratios* está disponível na maioria dos programas que calculam metanálise. Alguns métodos incluem o teste “*log-rank*” e regressão de Cox. Também é possível calcular o *hazard ratio* indiretamente, por manipulação do valor de qui-quadrado, valor de *p* ou variância⁸⁸. A escolha particular do método vai depender de quais informações estão disponíveis em um estudo particular.

Mesmo se os dados necessários para estimar o *hazard ratio* ou eles próprios forem fornecidos, o *hazard ratio* como medida de efeito para o cálculo da metanálise é preferível em metanálise de dados individuais de pacientes. Neste tipo de revisão, obtêm-se os dados da revisão diretamente dos bancos de dados dos estudos primários, o que permite uma informação mais acurada.

B) Medidas de efeito de desfechos ordinais

Se comportam, na maioria das vezes, como dados contínuos, na medida em que se aumenta o número de categorias e, portanto, são apresentados com as mesmas medidas de efeito usadas para dados contínuos (diferença de médias e diferença de médias padronizadas).

Entretanto, dados ordinais também podem assumir padrão de dados dicotômicos, quando se combinam as categorias adjacentes e se estabelece um ponto de corte. Nestes casos, utilizam-se como medidas de desfechos as mesmas medidas utilizadas para os dados dicotômicos [risco relativo (RR), *odds ratio* (OR), diferenças dos riscos (RRR), *hazard ratio* (HR)].

Apêndice G – Interpretando o gráfico de floresta

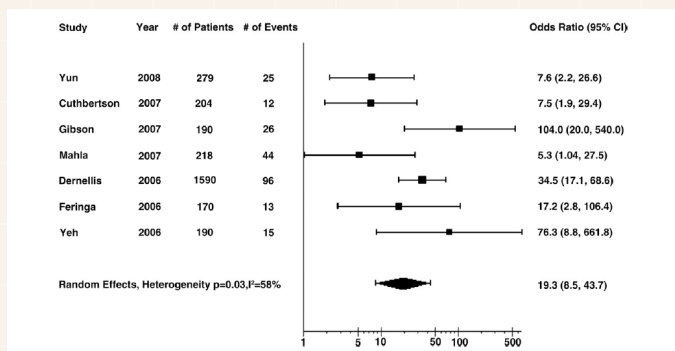
O gráfico de floresta ou *Forest plot*, como é conhecido em inglês, é um gráfico bastante útil, que apresenta diversas informações importantes, para interpretação dos resultados de uma metanálise. O gráfico de floresta apresenta as medidas de associação, com seus respectivos intervalos de confiança, para os estudos individuais e para a metanálise, tanto em texto, quanto em representação gráfica.

Cada estudo é apresentado em linhas horizontais, e pode ser identificado por nome do autor e data de publicação ou pelo acrônimo do estudo. Ao lado da identificação do estudo estão os dados de incidência do desfecho (número de caso sobre o total de pacientes no grupo intervenção e controle). A medida de associação entre a exposição e desfecho nos estudos individuais é apresentada em forma de um quadrado, com linhas em ambos os lados horizontais do quadrado, que representam o intervalo de confiança da estimativa (geralmente com 95% de nível de confiança). A área do quadrado representa o peso do estudo na metanálise e quanto maior esta área, maior é o número de eventos.

A metanálise é o diamante negro apresentado na porção inferior do gráfico e representa a medida sumária das medidas de associação.

Para interpretação dos resultados, estudos que concentram o quadrado preto à esquerda da linha vertical do gráfico indicam que a exposição diminui a ocorrência de eventos (exposição protetora) e os estudos que concentram seus quadrados pretos à direita da linha vertical indicam que a exposição aumenta a ocorrência de eventos (exposição maléfica). O mesmo se aplica para a interpretação da metanálise (diamante). A linha sólida vertical que divide os lados do gráfico é a linha da ausência de efeito. Se o diamante ou as linhas que representam os intervalos de confiança de cada estudo individual estiverem em cima da linha vertical da ausência de associação, a interpretação é que não existem diferenças estatisticamente significativas entre ser exposto ou não. A Figura G.1 apresenta um exemplo de gráfico de funil.

Figura G.1 – Odds Ratio ajustado para BNP/NT-proBNP pré-operatório na previsão de eventos cardiovasculares em 30 dias após a cirurgia



Fonte: (KARTHIKEYAN, G., 2009)⁸⁹.



Anexo A – Strengthening the Reporting of OBservational studies in Epidemiology (STROBE)

Itens essenciais que devem ser descritos em estudos observacionais, segundo a declaração Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) 2007 e traduzido por Malta M.; Cardoso, L. O.; Bastos, F. I.; Magnanini, M. M. F.; Silva, C. M. F. P.; 2010:

Item	Nº	Recomendação
Título e Resumo	1	Indique o desenho do estudo no título ou no resumo, com termo comumente utilizado.
Introdução		
Contexto/Justificativa	2	Detalhe o referencial teórico e as razões para executar a pesquisa.
Objetivos	3	Descreva os objetivos específicos, incluindo quaisquer hipóteses pré-existentes.
Métodos		
Desenho do estudo	4	Apresente, no início do artigo, os elementos-chave relativos ao desenho do estudo.
Contexto (<i>setting</i>)	5	Descreva o contexto, locais e datas relevantes, incluindo os períodos de recrutamento, exposição, acompanhamento (<i>follow-up</i>) e coleta de dados.
Participantes	6	Estudos de Coorte: Apresente os critérios de elegibilidade, fontes e métodos de seleção dos participantes. Descreva os métodos de acompanhamento. Estudos de Caso-Controlle: Apresente os critérios de elegibilidade, as fontes e o critério diagnóstico para identificação dos casos e os métodos de seleção dos controles. Descreva a justificativa para a eleição dos casos e controles. Estudo Seccional: Apresente os critérios de elegibilidade, as fontes e os métodos de seleção dos participantes. Estudos de Coorte: Para os estudos pareados, apresente os critérios de pareamento e o número de expostos e não expostos. Estudos de Caso-Controlle: Para os estudos pareados, apresente os critérios de pareamento e o número de controles para cada caso.
Variáveis	7	Defina claramente todos os desfechos, exposições, preditores, confundidores em potencial e modificadores de efeito. Quando necessário, apresente os critérios diagnósticos.
Fontes de dados/ Mensuração	8	Para cada variável de interesse, forneça a fonte dos dados e os detalhes dos métodos utilizados na avaliação (mensuração). Quando existir mais de um grupo, descreva a comparabilidade dos métodos de avaliação.
Viés	9	Especifique todas as medidas adotadas para evitar potenciais fontes de viés.
Tamanho do estudo	10	Explique como se determinou o tamanho amostral.
Variáveis quantitativas	11	Explique como foram tratadas as variáveis quantitativas na análise. Se aplicável, descreva as categorizações que foram adotadas e porque.

Continua

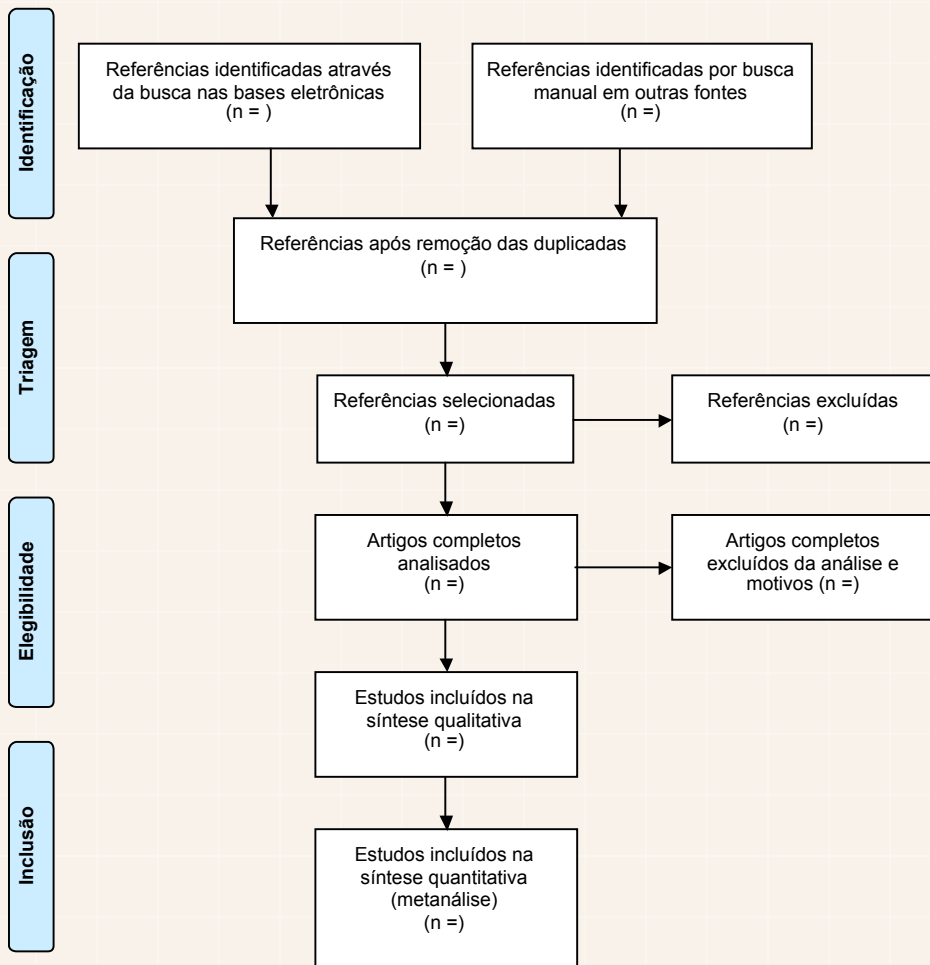
Continuação

Métodos estatísticos	12	Descreva todos os métodos estatísticos, incluindo aqueles usados para controle de confundimento. Descreva todos os métodos utilizados para examinar subgrupos e interações. Explique como foram tratados os dados faltantes (“missing data”). Estudos de Coorte: Se aplicável, explique como as perdas de acompanhamento foram tratadas. Estudos de Caso-Control: Se aplicável, explique como foi realizado o pareamento dos casos e controles. Estudos Seccionais: Se aplicável, descreva os métodos utilizados para considerar a estratégia de amostragem. Descreva qualquer análise de sensibilidade.
Resultados		
Participantes	13 ^a	Descreva o número de participantes em cada etapa do estudo (ex: número de participantes potencialmente elegíveis, examinados de acordo com critérios de elegibilidade, elegíveis de fato, incluídos no estudo, que terminaram o acompanhamento e efetivamente analisados).
Dados descritivos	14 ^a	Descreva as características dos participantes (por ex.: demográficas, clínicas e sociais) e as informações sobre exposições e confundidores em potencial. Indique o número de participantes com dados faltantes para cada variável de interesse. Estudos de Coorte: Apresente o período de acompanhamento (por ex.: média e tempo total).
Desfecho	15 ^a	Estudos de Coorte: Descreva o número de eventos-desfecho ou as medidas-resumo ao longo do tempo. Estudos de Caso-Control: Descreva o número de indivíduos em cada categoria de exposição ou apresente medidas-resumo de exposição. Estudos Seccionais: Descreva o número de eventos-desfecho ou apresente as medidas-resumo.
Resultados principais	16	Descreva as medidas de associação não ajustadas e, se aplicável, as estimativas ajustadas por variáveis confundidoras, assim como sua precisão (ex: intervalos de confiança). Deixe claro quais foram os confundidores utilizados no ajuste e porque foram incluídos. Quando variáveis contínuas forem categorizadas, informe os pontos de corte utilizados. Se pertinente, considere transformar as medidas de risco relativo em termos de risco absoluto, para um período de tempo relevante.
Outras análises	17	Descreva outras análises que tenham sido realizadas. Por ex.: análises de subgrupos, interação, sensibilidade.
Discussão		
Resultados principais	18	Resuma os principais achados relacionando-os aos objetivos do estudo.
Limitações	19	Apresente as limitações do estudo, levando em consideração fontes potenciais de viés ou imprecisão. Discuta a magnitude e direção de vieses em potencial.
Interpretação	20	Apresente uma interpretação cautelosa dos resultados, considerando os objetivos, as limitações, a multiplicidade das análises, os resultados de estudos semelhantes e outras evidências relevantes.
Generalização	21	Discuta a generalização (validade externa) dos resultados.
Outras Informações		
Financiamento	22	Especifique a fonte de financiamento do estudo e o papel dos financiadores. Se aplicável, apresente tais informações para o estudo original no qual o artigo é baseado.

Descreva essas informações separadamente para casos e controles em estudos de caso-controle e para grupos de expostos e não expostos, em estudos de coorte ou estudos seccionais.

Nota: Documentos mais detalhados discutem de forma mais aprofundada cada item do *checklist*, além de apresentarem o referencial teórico no qual essa lista se baseia e exemplos de descrições adequadas de cada item. O *checklist* do STROBE é mais adequadamente utilizado em conjunto com esses artigos (disponíveis gratuitamente no site das revistas PLoS Medicine⁹⁰, Annals of Internal Medicine⁹¹ e Epidemiology⁹²). No *website* da iniciativa STROBE⁹³ estão disponíveis versões separadas de *checklist* para estudos de Coorte, Caso-Controle ou Seccionais. Reproduzida de von Elm, E.; Altman, D. G.; Egger, M.; Pocock, S. J.; Gøtzsche, P. C.; Vandenbroucke, J. P. Declaração STROBE: Diretrizes para a comunicação de estudos observacionais (material suplementar na internet). Malta, M.; Cardoso, L. O.; tradutores. In: Malta, M.; Cardoso, L. O.; Bastos, F. I.; Magnanini, M. M. F.; Silva, C. M. F. P. Iniciativa STROBE: subsídios para a comunicação de estudos observacionais¹³.

/// Anexo B – Fluxo de seleção dos artigos da revisão sistemática



Fonte: Ministério da Saúde, 2012.

// Anexo C – Newcastle-Ottawa: Escala de avaliação de qualidade

Tradução simples (não validada):

105

NEWCASTLE – OTTAWA: Escala de Acesso de Qualidade

1. Para estudos de Coortes

Nota: Um estudo pode ser premiado com no máximo uma estrela (*) para cada item numerado dentro das categorias **Seleção e Desfecho**. Um máximo de duas estrelas pode ser dado para cada item da categoria **Comparabilidade**.

Seleção

1) Representatividade da Coorte exposta.

- a) verdadeiramente representativa da média _____ (descrever) na comunidade *
- b) um pouco representativa da média _____ na comunidade *
- c) grupo selecionado de usuários, por exemplo, enfermeiros, voluntários
- d) não há descrição de derivação da coorte

2) Seleção da Coorte não exposta.

- a) selecionada da mesma comunidade que a coorte exposta *
- b) selecionada a partir de uma fonte diferente
- c) não há descrição de derivação da coorte não exposta

3) Determinação da Exposição.

- a) registro seguro (p. ex., registros cirúrgicos) *
- b) entrevistas estruturadas *
- c) auto-relato escrito
- d) sem descrição

4) Demonstração que o desfecho de interesse não estava presente no início do estudo.

- a) sim *
- b) não

Comparabilidade

1) Comparabilidade da coorte baseada no desenho e análise.

- a) controles do estudo para _____ (selecione o fator mais importante) *
- b) controle do estudo para qualquer fator adicional * (este critério pode ser modificado para indicar controle específico para um segundo fator importante)

Desfecho

1) Determinação do desfecho.

- a) avaliação cega independente *
- b) registro acoplado *
- c) auto-relato
- d) sem descrição

2) O seguimento foi suficiente para a ocorrência dos desfechos?

- a) sim (escolha um adequado período de seguimento para o desfecho de interesse) *
- b) não

3) Adequação de acompanhamento da coortes.

- a) seguimento completo – todos os indivíduos *
- b) perdas de seguimento com improvável introdução de viés – pequeno número perdido – > ____ % (selecione um adequado %) seguimento, ou descrição fornecida daqueles perdidos)*
- c) taxa de seguimento < ____% (selecione um adequado %) e sem descrição das perdas
- d) nenhuma declaração

2. Para estudos de Caso-Controlle

Nota: Um estudo pode ser premiado com no máximo uma estrela (*) para cada item numerado dentro das categorias **Seleção e Exposição**. Um máximo de duas estrelas pode ser dado para cada item da categoria **Comparabilidade**.

Seleção

1) A definição de caso é adequada?

- a) sim, com validação independente *
- b) sim, por exemplo, registro de pareamento ou baseado em auto-relato
- c) sem descrição

2) Representatividade dos casos.

- a) consecutivos ou obviamente representativo de série de casos *
- b) potencial para viés de seleção ou não indicado

3) Seleção dos Controles.

- a) controles provenientes da comunidade *
- b) controles provenientes do hospital
- c) sem descrição

4) Definição dos Controles.

- a) sem história de doenças (desfecho) *
- b) sem descrição da fonte

Comparabilidade

1) Comparabilidade de casos e controles baseados no desenho ou na análise.

- a) controles do estudo para _____ (selecione o fator mais importante) *
- b) controles do estudo para qualquer fator adicional * (este critério pode ser modificado para indicar controle específico para um segundo fator importante)

Exposição

1) Determinação da exposição.

- a) registro seguro (p. ex., registros cirúrgicos) *
- b) entrevista estruturada onde o *status* caso/controle é “cego/mascarado” *
- c) entrevista aberta para o *status* caso/controle
- d) auto relatório escrito ou registro médico apenas
- e) sem descrição

2) Mesmo método de determinação para casos e controles.

- a) sim *
- b) não

3) Taxa de não-resposta.

- a) mesma taxa para ambos os grupos *
- b) não-respondedores descritos
- c) taxa diferente e sem designação

Instrumento original:

NEWCASTLE-OTTAWA QUALITY ASSESSMENT SCALE

1. For cohort studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort.

- a) truly representative of the average _____ (describe) in the community*
- b) somewhat representative of the average _____ in the community *
- c) selected group of users e.g. nurses, volunteers
- d) no description of the derivation of the cohort

2) Selection of the non exposed cohort.

- a) drawn from the same community as the exposed cohort *
- b) drawn from a different source
- c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure.

- a) secure record (e.g. surgical records) *
- b) structured interview *
- c) written self report
- d) no description

4) Demonstration that outcome of interest was not present at start of study.

- a) yes *
- b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis.

- a) study controls for _____ (select the most important factor) *
- b) study controls for any additional factor * (this criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome.

- a) independent blind assessment *
- b) record linkage *
- c) self report
- d) no description

2) Was follow-up long enough for outcomes to occur.

- a) yes (select an adequate follow up period for outcome of interest) *
- b) no

3) Adequacy of follow up of cohorts.

- a) complete follow up – all subjects accounted for *
- b) subjects lost to follow up unlikely to introduce bias – small number lost – > ____ % (select an adequate %) follow up, or description provided of those lost) *
- c) follow up rate < ____% (select an adequate %) and no description of those lost
- d) no statement

2. For case-control studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?

- a) yes, with independent validation *
- b) yes, e.g. record linkage or based on self reports
- c) no description

2) Representativeness of the cases.

- a) consecutive or obviously representative series of cases *
- b) potential for selection biases or not stated

3) Selection of Controls.

- a) community controls *
- b) hospital controls
- c) no description

4) Definition of Controls.

- a) no history of disease (endpoint) *
- b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis.

- a) study controls for _____ (select the most important factor.) *
- b) study controls for any additional factor * (this criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure.

- a) secure record (eg surgical records) *
- b) structured interview where blind to case/control status *
- c) interview not blinded to case/control status
- d) written self report or medical record only
- e) no description

2) Same method of ascertainment for cases and controls.

- a) yes *
- b) no

3) Non-Response rate.

- a) same rate for both groups *
- b) non respondents described
- c) rate different and no designation

Fonte: (Von, E., 2013)⁴⁸.

/// Anexo D – *Checklist Downs and Black* adaptado para avaliação da qualidade metodológica

Tradução simples (não validada):

Domínio: Relato

1. A hipótese/objetivo/estudo está claramente descrita?

Resposta	Escore
Sim	1
Não	0

2. Os desfechos principais a serem avaliados estão claramente descritos na seção Introdução ou Métodos?

Se os desfechos principais estiverem mencionados pela primeira vez na seção Resultados, a resposta para a questão deve ser “não”.

Resposta	Escore
Sim	1
Não	0

3. As características dos pacientes incluídos estão claramente descritas?

Em estudos de coorte os critérios de inclusão e exclusão devem estar descritos. Em estudos de caso-controle, deve ser apresentada a definição de caso e a fonte de controles.

Resposta	Escore
Sim	1
Não	0

4. As intervenções (exposições) estão claramente descritas?

A exposição/intervenção e o controle (não exposição) devem estar claramente descritos.

Resposta	Escore
Sim	1
Não	0

5. A distribuição dos principais fatores de confusão nos grupos de comparação está claramente descrita?

Considere “sim” quando uma lista dos principais fatores de confusão for disponibilizada.

Resposta	Escore
Sim	2
Parcialmente	1
Não	0

6. Os principais resultados do estudo estão claramente descritos?

Dados “simples” (incluindo denominadores e numeradores) devem ser relatados para todos os desfechos principais, de modo que o leitor possa conferir as principais análises e conclusões (esta questão não abrange os testes estatísticos que são considerados abaixo).

Resposta	Escore
Sim	1
Não	0

7. O estudo apresentou as estimativas de dispersão e variabilidade para os principais resultados?

Para dados com distribuição não paramétrica deve ser relatado o intervalo inter-quartil dos resultados. Para dados com distribuição normal devem ser reportados o erro padrão, desvio padrão ou intervalo de confiança. Se a distribuição dos dados não estiver descrita, deve-se assumir que as estimativas utilizadas foram adequadas e a questão deve ser respondida “sim”.

Resposta	Escore
Sim	1
Não	0

8. Todos os eventos adversos importantes foram reportados?

Questão deve ser respondida “sim” se o estudo demonstrar que foram feitos esforços abrangentes para avaliar eventos adversos. (Uma lista de potenciais eventos adversos foi fornecida).

Resposta	Escore
Sim	1
Não	0

9. As características dos pacientes com perda de seguimento foram descritas?

Questão deve ser respondida “sim” caso não houver perdas de seguimento ou se as perdas tiverem sido pequenas a ponto de não comprometerem as conclusões do estudo. Deve ser respondida “não” se o estudo não informar o número de pacientes com perda de seguimento.

Resposta	Escore
Sim	1
Não	0

10. Foram apresentados os valores exatos de probabilidade (por exemplo, 0,035 em vez de $<0,05$) para os desfechos principais (exceto quando a probabilidade for menos que 0,001)?

Resposta	Escore
Sim	1
Não	0

Domínio: Validade Externa

Todos os seguintes critérios abordam a representatividade dos resultados e generalização para a população a partir da qual foram obtidos os participantes do estudo.

11. Os participantes convidados para o estudo são representativos de toda a população de onde foram recrutados?

O estudo deve identificar a população de origem dos participantes e descrever como os participantes foram selecionados. Os participantes são representativos se compreenderem toda a população de origem, uma amostra não selecionada de pacientes consecutivos, ou uma amostra aleatória. A amostragem aleatória só é possível onde existir uma lista de todos os membros da população relevante. Quando o estudo não relatar a proporção da população de origem a partir do qual os participantes foram selecionados, a questão deve ser respondida como “indeterminado”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

12. Os indivíduos incluídos no estudo são representativos de toda a população de onde foram recrutados?

A proporção dos participantes convidados que concordou em participar deve ser indicada. A amostra é considerada representativa se for demonstrado que a distribuição dos principais fatores de confusão é a mesma na amostra do estudo e na população de origem.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

13. A equipe de assistência à saúde, o local e as instalações onde os participantes do estudo foram tratados (expostos ou recrutados) são representativos em relação ao que a maioria dos pacientes/população recebe?

Para a pergunta deve ser respondida como “sim” se o estudo demonstrar que a intervenção foi representativa daquela em uso na população fonte. A questão deve ser respondida como “não” se, por exemplo, a intervenção foi realizada em um centro especializado, não representativo dos hospitais onde a maioria da população receberia assistência.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

Domínio: Viés

14. Foram feitos esforços para manter pacientes sem conhecimento sobre a exposição (cegamento – não aplicável para a maioria dos estudos observacionais)?

Para a pergunta deve ser respondida como “sim” nos estudos em que os pacientes não sabem qual intervenção receberam.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

15. Foram feitos esforços para manter avaliadores de desfechos sem conhecimento sobre a exposição?

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

16. Algum dos resultados foi baseado em análise exploratória *post hoc*? Se sim, está claramente descrito?

Qualquer análise que não tenha sido planejada no início do estudo deve ser claramente indicada. Se nenhuma análise de subgrupo *post hoc* tiver sido conduzida, responda “sim”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

17. Em estudos clínicos randomizados e estudos de coorte, as análises foram ajustadas para diferentes tempos de seguimento? Ou em estudos de caso-controle, o tempo entre a exposição e desfecho foi o mesmo para casos e controles?

Se o tempo de seguimento for o mesmo para todos os participantes, responda “sim”. Se diferentes tempos de seguimentos tiverem sido considerados, mas estiverem ajustados como, por exemplo, em uma análise de sobrevida, responda “sim”. Se os tempos de seguimento tiverem sido ignorados, responda “não”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

18. Os testes estatísticos utilizados nas análises primárias foram adequados?

As análises estatísticas devem ser realizadas considerando testes adequados. Por exemplo, métodos não paramétricos devem ser usados para pequenas amostras. Quando poucas análises estatísticas tiverem sido realizadas, e não houver evidência de viés, a questão deve ser respondida “sim”. Se a distribuição dos dados não estiver descrita (paramétrico ou não), deve-se supor que as estimativas utilizadas foram adequadas e a questão deve ser respondida “sim”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

19. Houve aderência à intervenção (se aplicável)?

Quando não houver adesão ao tratamento alocado/exposição ou quando houver contaminação de um grupo, a questão deve ser respondida “não”. Quando houver viés de aferição diferencial (ou outro viés) que influencie a estimativa de associação no sentido do efeito nulo (sem diferença entre os grupos), a questão deve ser respondida “sim”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

20. A medida de desfecho primário foi acurada (válida e confiável)?

Para estudos onde as medidas de desfecho estiverem claramente descritas, a questão deve ser respondida “sim”. Para estudos que referenciem outras publicações ou demonstrarem que as medidas são acuradas, a questão deve ser respondida como “sim”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

Domínio: Variável de confusão (viés de seleção)

21. Os pacientes dos grupos expostos e não expostos (para ensaios clínicos e estudos de coorte) foram recrutados da mesma população? Ou os casos e controles (em estudos de caso-controle) foram recrutados da mesma população?

Por exemplo, para todos os grupos de comparação os pacientes devem ter sido selecionados a partir do mesmo hospital. A questão deve ser respondida como “indeterminada” quando não houver informação sobre a fonte de pacientes.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

22. Os pacientes dos grupos expostos e não expostos (para ensaios clínicos e estudos de coorte) foram recrutados no mesmo período? Ou os casos e controles (em estudos de caso-controle) foram recrutados no mesmo período?

Para um estudo que não especifica o período de tempo durante o qual os pacientes foram recrutados, a questão deve ser respondida como “indeterminada”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

23. Os participantes foram randomizados para os grupos de intervenção?

Para estudos que claramente descreverem que houve randomização, a questão deve ser respondida como “sim”, exceto se o método de randomização utilizado não assegurar alocação aleatória. Por exemplo, se o estudo utilizou alternância de participantes, a questão deve ser respondida como “não”, pois a alocação é previsível. Para estudos observacionais responder “não”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

24. A alocação aleatória da intervenção foi sigilosa para pacientes e equipe de assistência à saúde até a completa inclusão do paciente no estudo?

Para todos os estudos não randomizados a resposta deve ser “não”. Se a alocação foi sigilosa para pacientes, mas não para a equipe, a resposta deve ser “não”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

25. As análises ajustadas foram adequadas em relação às variáveis de confusão a partir da qual as principais conclusões foram tomadas?

Esta questão deve ser respondida “não” para ensaios clínicos randomizados se: as principais conclusões do estudo tiverem sido baseadas em análises por protocolo em vez de intenção de tratar; a distribuição das variáveis de confusão entre os grupos não estiver descrita; ou a distribuição de fatores de confusão conhecidos estiver desbalanceada entre os grupos e esta diferença não estiver considerada nas análises. Em estudos não randomizados, se o efeito dos principais fatores de confusão não foi investigado ou o fator de confusão foi demonstrado, mas nenhum ajuste foi feito na análise final, a questão deve ser respondida como “não”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

26. As perdas de seguimento foram consideradas?

Se o número de pacientes que apresentaram perda de seguimento não estiver reportado, a questão deve ser respondida como “indeterminada”. Se as perdas tiverem sido pequenas a ponto de não comprometerem os resultados principais, a questão deve ser respondida “sim”.

Resposta	Escore
Sim	1
Não	0
Indeterminado	0

Domínio: Poder

27. O estudo tem poder suficiente para detectar uma diferença clinicamente importante entre os grupos quando a probabilidade desta diferença ter ocorrido ao acaso for menor que 5%?

Resposta	Escore
Sim	1
Não	0

Instrumento original:

Domain: Reporting

1. Is the hypothesis/aim/objective of the study clearly described?

Answer	Score
Yes	1
No	0

2. Are the main outcomes to be measured clearly described in the Introduction or Methods section?

If the main outcomes are first mentioned in the Results section, the question should be answered no.

Answer	Score
Yes	1
No	0

3. Are the characteristics of the patients included in the study clearly described?

In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given.

Answer	Score
Yes	1
No	0

4. Are the interventions of interest clearly described?

Treatments and placebo (where relevant) that are to be compared should be clearly described.

Answer	Score
Yes	1
No	0

5. Are the distributions of principal confounders in each group of subjects to be compared clearly described?

A list of principal confounders is provided.

Answer	Score
Yes	2
Partially	1
No	0

6. Are the main findings of the study clearly described?

Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical tests which are considered below).

Answer	Score
Yes	1
No	0

7. Does the study provide estimates of the random variability in the data for the main outcomes?

In non normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.

Answer	Score
Yes	1
No	0

8. Have all important adverse events that may be a consequence of the intervention been reported?

This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).

Answer	Score
Yes	1
No	0

9. Have the characteristics of patients lost to follow-up been described?

This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does not report the number of patients lost to follow-up.

Answer	Score
Yes	1
No	0

10. Have actual probability values been reported (e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?

Answer	Score
Yes	1
No	0

Domain: External validity

All the following criteria attempt to address the representativeness of the findings of the study and whether they may be generalised to the population from which the study subjects were derived.

11. Were the subjects asked to participate in the study representative of the entire population from which they were recruited?

The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample. Random sampling is only feasible where a list of all members of the relevant population exists. Where a study does not report the proportion of the source

population from which the patients are derived, the question should be answered as unable to determine.

Answer	Score
Yes	1
No	0
Unable to determine	0

12. Were those subjects who were prepared to participate representative of the entire population from which they were recruited?

The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.

Answer	Score
Yes	1
No	0
Unable to determine	0

13. Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?

For the question to be answered yes the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.

Domain: Bias

14. Was an attempt made to blind study subjects to the intervention they have received?

For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.

Answer	Score
Yes	1
No	0
Unable to determine	0

15. Was an attempt made to blind those measuring the main outcomes of the intervention?

Answer	Score
Yes	1
No	0
Unable to determine	0

16. If any of the results of the study were based on “data dredging”, was this made clear?

Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.

Answer	Score
Yes	1
No	0
Unable to determine	0

17. In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls?

Where follow-up was the same for all study patients the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.

Answer	Score
Yes	1
No	0
Unable to determine	0

18. Were the statistical tests used to assess the main outcomes appropriate?

The statistical techniques used must be appropriate to the data. For example nonparametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.

Answer	Score
Yes	1
No	0
Unable to determine	0

20. Were the main outcome measures used accurate (valid and reliable)?

For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.

Answer	Score
Yes	1
No	0
Unable to determine	0

Domain: Confounding (selection bias)

21. Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?

For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and casecontrol studies where there is no information concerning the source of patients included in the study.

Answer	Score
Yes	1
No	0
Unable to determine	0

22. Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?

For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.

Answer	Score
Yes	1
No	0
Unable to determine	0

23. Were study subjects randomised to intervention groups?

Studies which state that subjects were randomised should be answered yes except where method of randomisation would not ensure random allocation. For example alternate allocation would score no because it is predictable.

Answer	Score
Yes	1
No	0
Unable to determine	0

24. Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?

All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.

Answer	Score
Yes	1
No	0
Unable to determine	0

25. Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?

This question should be answered no for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In nonrandomized studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.

Answer	Score
Yes	1
No	0
Unable to determine	0

26. Were losses of patients to follow-up taken into account?

If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.

Answer	Score
Yes	1
No	0
Unable to determine	0

27. Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?

Answer	Score
Yes	1
No	0

Fonte: (Downs e cols., 1998, adaptado)⁴⁷.

/// Anexo E – Diretriz de redação do manuscrito – Meta-analysis of observational studies in epidemiology (MOOSE)²

Tradução simples (não validada):

Um *checklist* para autores, editores e revisores de metanálises de estudos observacionais:

A introdução deve conter:

- definição do problema;
- hipótese;
- descrição do(s) desfecho(s) do estudo;
- tipo de exposição ou intervenção utilizada;
- tipo de desenhos de estudo utilizados; e
- população do estudo.

A estratégia de busca deve conter:

- qualificação dos pesquisadores (por exemplo, bibliotecários e pesquisadores);
- estratégia de busca, incluindo o período de tempo envolvido e palavras-chave;
- quais esforços foram realizados para incluir todos os estudos disponíveis, incluindo contato com os autores dos estudos primários;
- bases de dados e registros pesquisados;
- *software* de busca utilizado, nome e versão, incluindo recursos especiais utilizados (por exemplo, “explosão”);
- uso busca manual (por exemplo, listas de referências dos artigos obtidos);
- lista de citações recuperadas pela estratégia e busca e lista de citações excluídas, com suas justificativas;
- método para manejo de artigos publicados em outros idiomas além do Inglês;
- método para manejo de resumos e estudos não publicados; e
- descrição de qualquer contato com os autores.

A seção de métodos deve incluir:

- descrição da relevância ou adequação dos estudos agrupados para avaliar a hipótese a ser testada;
- justificativa para a seleção e codificação de dados (por exemplo, princípios clínicos ou conveniência);
- documentação de como os dados foram classificados e codificados (por exemplo, revisores independentes, “cegamento” e confiabilidade entre revisores);
- avaliação dos fatores de confusão (por exemplo, se a comparabilidade dos casos e controles foi apropriada);
- avaliação da qualidade do estudo, incluindo “cegamento” dos avaliadores/revisores;

- uso de análise estratificada ou de regressão considerando as possíveis variáveis preditoras de desfecho;
- avaliação da heterogeneidade;
- descrição dos métodos estatísticos (por exemplo, descrição completa de modelos de efeitos fixos ou aleatórios, considerando a justificativa para escolha do método e se o modelo leva em conta as variáveis preditoras, efeito de dose-resposta, ou meta-análise cumulativa) com detalhes suficientes para serem replicados; e
- se serão apresentados tabelas e gráficos.

A seção de resultados deve incluir:

- resumo gráfico das estimativas dos estudos primários e a estimativa global de efeito;
- um quadro com informações descritivas para cada estudo incluído;
- resultados dos testes de sensibilidade (por exemplo, análise de subgrupo); e
- indicação de incerteza estatística dos resultados.

A discussão deve incluir:

- avaliação quantitativa do risco de viés (por exemplo, viés de publicação);
- justificativa para a exclusão (por exemplo, a exclusão de estudos em idioma não inglês);
- avaliação da qualidade dos estudos incluídos; e
- a discussão também deve incluir a discussão de questões relacionadas ao risco de viés, incluindo viés de publicação, variáveis de confusão e outros vieses.

A conclusão deve incluir:

- consideração de explicações alternativas para os resultados observados;
- generalização das conclusões (ou seja, adequadas para os dados apresentados);
- considerações para futuras pesquisas; e
- divulgação da fonte de financiamento.

Instrumento original:

A checklist for authors, editors, and reviewers of meta-analyses of observational Studies².

Reporting of the background should include:

- problem definition;
- hypothesis statement;
- description of study outcome(s);
- type of exposure or intervention used;
- type of study designs used; and
- study population.

Reporting of the search strategy should include:

- qualifications of searchers (e.g., librarians and investigators);
- search strategy, including time period included in the synthesis and keywords;
- effort to include all available studies, including contact with authors;
- databases and registries searched;
- search software used, name and version, including special features used (e.g., explosion);
- use of hand searching (e.g., reference lists of obtained articles);
- list of citations located and those excluded, including justification;
- method of addressing articles published in languages other than English;
- method of handling abstracts and unpublished studies; and
- description of any contact with authors.

Reporting of methods should include:

- description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested;
- rationale for the selection and coding of data (e.g., sound clinical principles or convenience);
- documentation of how data were classified and coded (e.g., multiple raters, blinding, and inter-rater reliability);
- assessment of confounding (e.g., comparability of cases and controls in studies where appropriate);
- assessment of study quality, including blinding of quality assessors;
- stratification or regression on possible predictors of study results;
- assessment of heterogeneity;
- description of statistical methods (e.g., complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated; and
- provision of appropriate tables and graphics.

Reporting of results should include:

- graphical summary of individual study estimates and the overall estimate of effect;
- a table giving descriptive information for each study included;
- results of sensitivity testing (e.g., subgroup analysis); and
- indication of statistical uncertainty of findings.

Reporting of discussion should include:

- quantitative assessment of bias (e.g., publication bias);
- justification for exclusion (e.g., exclusion of non-English-language citations);
- assessment of quality of included studies; and
- the discussion should also include discussion of issues related to bias, including publication bias, confounding, and quality.

Reporting of conclusions should include:

- consideration of alternative explanations for observed results;
- generalizability of the conclusions (i.e., appropriate to the data presented);
- guidelines for future research; and
- disclosure of funding source.

Esta obra foi impressa em papel couché fosco 240 g/m² (capa) e papel off set 90 g/m² (miolo) pela Nome da Gráfica, em agosto de 2014. A Editora do Ministério da Saúde foi responsável pela normalização (OS 2014/0529).

ISBN 978-85-334-2171-4



9 788533 421714