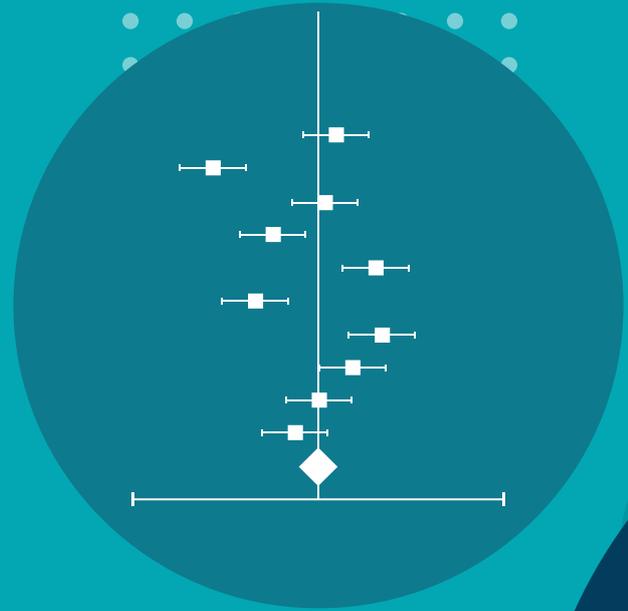


MINISTÉRIO DA SAÚDE



DIRETRIZES METODOLÓGICAS

ELABORAÇÃO DE
REVISÃO SISTEMÁTICA E
META-ANÁLISE DE ENSAIOS
CLÍNICOS RANDOMIZADOS



MINISTÉRIO DA SAÚDE
Secretaria de Ciência, Tecnologia e Insumos Estratégicos
Departamento de Gestão e Incorporação de Tecnologias em Saúde

DIRETRIZES METODOLÓGICAS:

ELABORAÇÃO DE
REVISÃO SISTEMÁTICA E
META-ANÁLISE DE ENSAIOS
CLÍNICOS RANDOMIZADOS



Brasília - DF
2020

2021 Ministério da Saúde.



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: <http://bvsmms.saude.gov.br>.

Tiragem: 1ª edição - 2021 - versão eletrônica

Elaboração, distribuição e informações:

MINISTÉRIO DA SAÚDE

Secretaria de Ciência, Tecnologia,
Inovação e Insumos Estratégicos em Saúde
Esplanada dos Ministérios,
Bloco G, Ed. Sede, Sobreloja
CEP: 70058-900 - Brasília/DF
Tel.: (61) 3315-6248
Site: www.saude.gov.br
E-mail: rebrats@saude.gov.br

Elaboração - Haoc/Dats:

Cleusa Pinheiro Ferri
Matheus Oliveira de Almeida
Thais Montezuma

Coordenação-Executiva:

Ana Paula Neves M. de Pinho
Lígia Fonseca Spinel
Nídia Cristina de Souza

Revisão técnica:

Betânia Ferreira Leite
Dalila Fernandes Gomes

Gabriela Viela de Brito
Gustavo Laine Araújo de Oliveira
Paulo Henrique Ribeiro Fernandes Almeida
Rachel Riera
Rafael Pacheco
Tais Galvão
Vania Eloisa

Supervisão-Geral:

Clementina Corah Lucas Prado
Gustavo Luís Meffe Andreoli
Hélio Angotti Neto
Vania Cristina Canuto Santos

Capa, projeto gráfico e diagramação:

Laura Camilo (L7Design)

Normalização:

Delano de Aquino Silva - Editora MS/CGDI

Revisão:

Khamila Silva - Editora MS/CGDI
Tamires Felipe Alcântara - Editora MS/CGDI

Ficha Catalográfica

Brasil. Ministério da Saúde. Secretaria de Ciência, Tecnologia, Inovação e Insumos Estratégicos em Saúde. Departamento de Gestão e Incorporação de Tecnologias em Saúde.

Diretrizes metodológicas : elaboração de revisão sistemática e meta-análise de ensaios clínicos randomizados [recurso eletrônico] / Ministério da Saúde, Secretaria de Ciência, Tecnologia, Inovação e Insumos Estratégicos em Saúde, Departamento de Gestão e Incorporação de Tecnologias em Saúde. - Brasília : Ministério da Saúde, 2021.

93 p. : il.

Modo de acesso: World Wide Web:

http://bvsmms.saude.gov.br/bvs/publicacoes/diretrizes_elaboracao_revisao_sistematica_meta-analise.pdf

ISBN 978-65-5993-021-0

1. Avaliação de Tecnologias em Saúde. 2. Revisão Sistemática. 3. Tomada de Decisão Clínica. I. Título.

CDU 614:62

Catálogo na fonte - Coordenação-Geral de Documentação e Informação - Editora MS - OS 2021/0022

Título para indexação:

Methodological guideline: development of systematic review and meta-analysis of randomized clinical trials

Lista de siglas e abreviaturas

AMED	Allied and Complementary Medicine Database
AMSTAR-2	Assessing the Metodological Quality of Systematic Reviews
Capes	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CINHAL	Cumulative Index to Nursing and Allied Health Literature
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
CRD	Centre for Reviews and Dissemination
DeCS	Descritores em Ciências da Saúde
DM	Diferença de médias
DMP	Diferença de médias padronizada
ECR	Ensaio clínico randomizado
Fapesp	Fundação de Amparo à Pesquisa do Estado de São Paulo
FDA	Food and Drug Administration
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
ICTRP	International Clinical Trials Registry Platform
IC	Intervalo de confiança
ID	Número de identificação do estudo
ITT	<i>Intention-to-treat</i> [intenção de tratar]
Lilacs	Literatura Latino-Americana e do Caribe em Ciências da Saúde
MECIR	Methodological Expectations of Cochrane Intervention Reviews
MEDLINE	Medical Literature Analysis and Retrieval System Online
MeSH	Medical Subject Headings
NNT	<i>Number needed to treat</i> [número necessário para tratar]
OR	<i>Odds ratio</i> [razão de chance]
PEDro	Physiotherapy Evidence Database
PICO	População, Intervenção, Comparador e Desfecho (em inglês, <i>outcome</i>)
PRISMA	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i>
PRISMA-P	<i>Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols</i>
PROSPERO	International Prospective Register of Systematic Reviews
RD	<i>Risk difference</i> [diferença de risco]
ReBEC	Registro Brasileiro de Ensaio Clínicos
RevMan	Review Manager
RoB	Risk of Bias
ROBIS	Risk of Bias in Systematic Reviews
RR	<i>Risk ratio or relative risk</i>
RS	Revisão sistemática
SRDR	Systematic Review Data Repository
TOI	Tamanho ótimo da informação
TOR	Tamanho ótimo da revisão
WHO	World Health Organization

Lista de quadros

Quadro 1	Exemplos de perguntas de pesquisa para RS	16
Quadro 2	Exemplos de filtros para ECR	20
Quadro 3	Vocabulários controlados das principais bases de dados	21
Quadro 4	Operadores booleanos	23
Quadro 5	Exemplo de questões que podem ser incluídas no formulário de extração	27
Quadro 6	Exemplo de <i>codebook</i> (guia codificado)	29
Quadro 7	Cálculo do risco relativo, razão de chance, diferença de risco e número necessário para tratar, com uma tabela de desfecho dicotômico	31
Quadro 8	Critérios para julgamento geral do risco de viés de ECR para um desfecho específico	38
Quadro 9	Avaliação da qualidade da evidência	55

Lista de figuras

Figura 1	Exemplo de <i>forest plot</i> com desfecho contínuo.	42
Figura 2	Principais métodos estatísticos para meta-análise de desfechos contínuos e dicotômicos	45
Figura 3	Exemplo do gráfico do funil com linha do efeito nulo (linha cinza-claro) e linha do efeito sumário (linha cinza-escuro)	52
Figura 4	Exemplo do gráfico do funil com linhas de significância	53
Figura 5	Passo a passo de como deve ser realizada a análise da qualidade de evidência em revisões sistemáticas de ensaios clínicos randomizados	56
Figura 6	Fluxograma do processo de seleção dos estudos	60
Figura 7	Exemplo de quadro de caracterização dos ensaios clínicos randomizados incluídos	61
Figura 8	Exemplo do resultado da avaliação do risco de viés usando a ferramenta RoB 2.0	62
Figura 9	Exemplo de <i>forest plot</i> com adição do resultado da avaliação do risco de viés	63
Figura 10	Exemplo de quadro de resultados dos desfechos dos estudos incluídos por comparação	64
Figura 11	Exemplo de quadro-sumário dos resultados para avaliação da qualidade da evidência	66

Sumário

APRESENTAÇÃO	6
1 INTRODUÇÃO ÀS REVISÕES SISTEMÁTICAS	7
1.1 Definição de revisão sistemática	7
1.2 Definição de meta-análise	7
1.3 Importância das revisões sistemáticas	8
1.4 Limitações e desafios da elaboração de revisões sistemáticas	9
1.5 Recursos necessários	9
2 PLANEJAMENTO DA REVISÃO SISTEMÁTICA	11
2.1 Justificativa para realizar uma revisão sistemática	11
2.2 Elaboração e registro do protocolo	12
2.3 Ferramentas importantes para elaboração da revisão sistemática	13
3 CONDUÇÃO E RELATO DA REVISÃO SISTEMÁTICA	15
3.1 Título	15
3.2 Resumo estruturado	15
3.3 Introdução	15
3.4 Objetivos da revisão sistemática	16
3.5 Métodos	16
3.6 Resultados	60
3.7 Discussão	69
3.8 Conclusão	70
REFERÊNCIAS	71
ANEXOS	78
Anexo A - Tradução oficial do checklist PRISMA (www.prisma-statement.org)	78
Anexo B - Principais bases de dados e seus respectivos endereços eletrônicos	80
Anexo C - Passo a passo de construção de estratégia de busca nas principais bases de dados	81
Anexo D - Exemplos de estratégias de busca	85
Anexo E - Tradução não validada da ferramenta RoB-2 para avaliação do risco de viés de ensaios clínicos randomizados	92

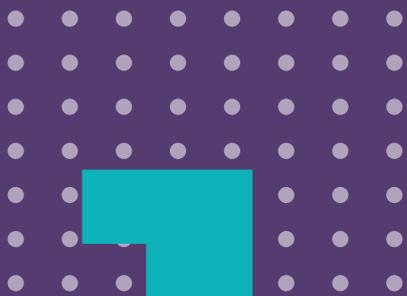
Apresentação

Revisões sistemáticas (RS) comumente produzem a melhor evidência disponível para avaliar a efetividade, a eficácia e a segurança de tecnologias, e são utilizadas para dar suporte às tomadas de decisões clínicas e nas políticas na saúde¹. A RS consiste em uma síntese crítica e reprodutível das melhores evidências disponíveis sobre uma questão específica, além de identificar lacunas sobre a pergunta de interesse que ainda precisam de respostas, como também identificar perguntas adicionais que possam surgir ao longo do desenvolvimento da RS. Por se tratar de método explícito e sistemático para identificar, selecionar e avaliar a qualidade de evidências, as RS são tipicamente estudos produzidos por uma metodologia confiável, rigorosa e reprodutível².

O objetivo deste documento é apresentar os principais conceitos relacionados ao planejamento, à condução, à interpretação e ao relato de RS, utilizando linguagem simples e acessível para profissionais da área da saúde, que estejam ou não familiarizados com esse delineamento de estudo. Essas diretrizes metodológicas são importantes para que revisões sistemáticas sejam conduzidas com maior qualidade, aumentando assim a confiança em seus achados.

Esta edição detalha, de forma pragmática, os processos que compõem uma RS e atualiza as ferramentas necessárias para sua elaboração, servindo como um guia prático para os autores de RS. O documento foi baseado em guias e diretrizes internacionais utilizados para planejar, conduzir e relatar RS de ensaios clínicos randomizados (ECR). Entre esses documentos internacionais se destacam:

- *The Cochrane Handbook for Systematic Review of Interventions*³.
- *Manual de Revisão Sistemática* do Joanna Briggs Institute⁴.
- *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)*⁵.
- *Methodological Expectations of Cochrane Intervention Reviews (MECIR)*⁶.



1

INTRODUÇÃO ÀS REVISÕES SISTEMÁTICAS



1.1 | Definição de revisão sistemática

A RS sintetiza os resultados de estudos primários que se enquadram nos critérios de elegibilidade para responder uma pergunta de pesquisa³. A busca por estudos é realizada de forma sistemática, por meio de várias fontes de dados, com o uso de uma estratégia de busca ampla e sensível⁷. A pergunta da pesquisa nas RS de ECR é estruturada no formato PICO (População, Intervenção, Comparador e Desfecho - em inglês, *outcome*). Esse tipo de revisão se baseia em métodos rigorosos, explícitos e reproduzíveis, para minimizar o risco de viés, entender as inconsistências dos resultados e então fornecer resultados confiáveis para a tomada de decisão ³.

1.2 | Definição de meta-análise

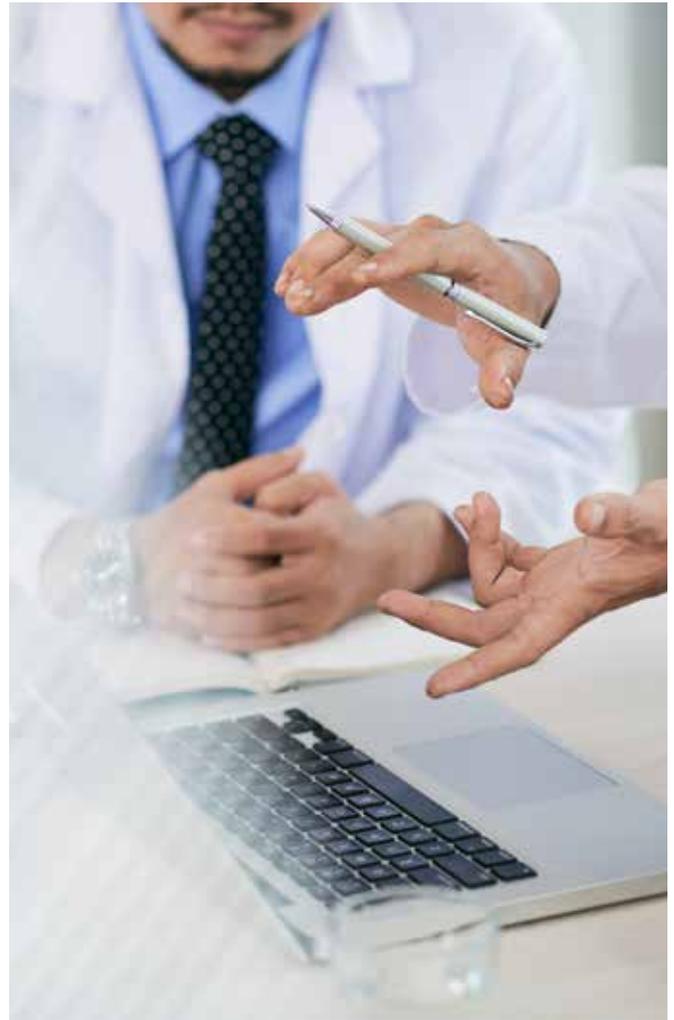
Meta-análise é um método estatístico que pode ser realizado em RS para combinar os resultados de dois ou mais estudos independentes, que podem testar ou não a mesma pergunta/hipótese, gerando uma única estimativa de efeito⁸. A meta-análise pode ser em pares/direta ou as meta-análises em rede, ou meta-análises de comparações múltiplas. Na presente diretriz, serão apresentados apenas os conceitos de meta-análise em pares. A vantagem da meta-análise é que esta pode apresentar uma estimativa mais precisa do tamanho do efeito, com considerável aumento do poder estatístico, o que é importante quando o poder do estudo primário é limitado devido ao pequeno tamanho da amostra⁹. Uma meta-análise pode, em alguns casos, produzir resultados conclusivos quando estudos individuais primários são inconclusivos.

Uma RS não necessariamente precisa apresentar uma meta-análise para sintetizar seus resultados. Ao contrário, em alguns casos não é apropriado que elas sejam realizadas, podendo até mesmo gerar conclusões errôneas. Nesse caso, sínteses qualitativas devem ser utilizadas¹⁰.

1.3 | Importância das revisões sistemáticas

A RS sintetiza os estudos primários da literatura sobre uma questão específica de pesquisa, auxiliando a tomada de decisão na saúde em torno da melhor conduta terapêutica ou profilática. Além disso, esse tipo de estudo ajuda a evitar desperdícios de pesquisa, garantindo que novas pesquisas primárias sejam realizadas com pleno conhecimento da evidência existente¹¹.

Um único estudo primário frequentemente não consegue detectar diferenças clinicamente relevantes de efeitos entre dois tratamentos, pois elaborar um ECR, com estimativas corretas para o cálculo do tamanho de amostra e poder estatístico adequados, exige, na maioria das vezes, um tamanho de amostra bastante significativo. Finalizar um ECR com um tamanho de amostra grande necessita de infraestrutura, recursos humanos especializados e recursos financeiros importantes, não disponíveis na maioria das vezes. De qualquer forma, RS com ou sem meta-análise são complementares aos estudos randomizados, por fornecerem ao profissional de saúde uma síntese do conhecimento atual disponível¹.



As RS também podem contribuir para aumentar a validade externa dos estudos, ou seja, contribuir para a generalização dos resultados. Em algumas situações, os achados de um estudo particular podem ser válidos apenas para uma população com as mesmas características das investigadas no estudo¹⁰.



Uma RS não necessariamente precisa apresentar uma meta-análise para sintetizar seus resultados.

1.4 | Limitações e desafios da elaboração de revisões sistemáticas

As RS possuem alguns desafios, como a gestão de grande volume de informações e extração de dados, o tempo para a realização do estudo e as possíveis limitações de recursos humanos e materiais disponíveis¹². As principais limitações que podem afetar a confiabilidade dos resultados de uma RS são o viés de publicação, além de dificuldades em combinar estudos que possam ter diferenças nas populações, nas intervenções, nos comparadores e na definição dos desfechos (diversidade clínica).

Assim, é necessário explorar criticamente a evidência apresentada pela RS. Interpretações enganosas podem ser evitadas se princípios básicos forem observados. Nessa diretriz, serão discutidas as etapas para realização de RS de ECR, com ênfase em princípios básicos para garantir a validade dos resultados.

1.5 | Recursos necessários

O conjunto de recursos, conhecimentos e habilidades da equipe envolvida é decisivo para a realização de uma RS de forma otimizada. O tempo é fator predominante e vai depender do tema e do escopo da revisão, assim como das habilidades e dos conhecimentos técnicos dos revisores. É preciso estimar o tempo que será necessário para a condução da RS e a disponibilidade que cada membro da equipe terá para trabalhar na revisão. A elaboração de um cronograma é importante, e este deve incluir todas as etapas, como treinamentos, reuniões, desenvolvimento do protocolo, busca e seleção dos estudos, extração e análise de dados etc.

É importante prever gastos, como a necessidade de contratar um tradutor, para tradução de estudos em idioma não dominado pela equipe, ou comprar estudos em texto completo, não disponíveis para os revisores. Assim, sugere-se a elaboração de um orçamento, contemplando todos os itens com seus respectivos valores.

Os recursos financeiros podem ser próprios ou podem ser financiados por entidades de agências de fomento à pesquisa (ex.: Conselho Nacional de Desenvolvimento Científico e Tecnológico [CNPq], Fundação de Amparo à Pesquisa do Estado de São Paulo [Fapesp] etc.), instituições responsáveis pela avaliação tecnológica (Ministério da Saúde) ou envolvidas na elaboração de diretrizes para prática clínica (sociedades brasileiras de especialidades clínicas).



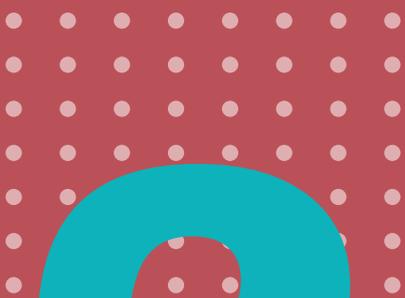
○ **Recursos humanos:** é importante estimar o número de pessoas que irão compor a equipe da revisão, em quais etapas, quem será o coordenador principal da revisão, a necessidade de consultoria de especialistas no assunto, além de apoio estatístico e bibliotecário (se necessário). É importante ter pelo menos três pesquisadores envolvidos para que seja possível minimizar os vieses e os erros em todos os estágios da RS⁷.

A equipe de pesquisa pode conter outros membros da sociedade envolvidos com o tema, os quais podem ser consultados em vários estágios da RS, como profissionais de saúde especialistas no tema pesquisado, representantes de pacientes, usuários de serviços e especialistas em métodos estatísticos⁷.

As habilidades e os conhecimentos são fundamentais para a garantia da qualidade metodológica da RS. Assim, conhecimentos aprofundados sobre métodos de pesquisa, epidemiologia, conhecimento clínico da questão de pesquisa, informática e domínio de língua inglesa são necessários. Além disso, a equipe deve saber utilizar as ferramentas necessárias à condução da revisão, como os gerenciadores de referência, bases de dados bibliográficas, softwares que auxiliem na condução ou no relato da revisão sistemática, programas estatísticos para meta-análise.

○ **Recursos materiais:** é importante estimar todos os recursos necessários para elaboração da RS, como computadores, acesso à internet, softwares de gerenciamento de referências (por exemplo, Mendeley, Endnote, Zotero), e softwares estatísticos (RevMan, R, Stata, Comprehensive Meta-Analysis, por exemplo), além de eventual aquisição de textos completos de publicações com acesso restrito.





2

PLANEJAMENTO DA REVISÃO SISTEMÁTICA



2.1 | Justificativa para realizar uma revisão sistemática

A condução de uma RS nova tem de ser justificada com uma pergunta de pesquisa definida. É importante verificar se ela já não foi respondida apropriadamente por uma RS prévia, bem como se há pesquisas primárias que potencialmente seriam incluídas pela nova RS.

Em algumas situações, pode existir uma RS já publicada, porém desatualizada. A viabilidade da atualização pode ser definida de acordo com os seguintes critérios: verificar se há pesquisas novas e relevantes e analisar a qualidade das revisões previamente publicadas. Outro ponto importante é confirmar a adequação do método de análise do risco de viés dos estudos individuais e a avaliação da qualidade da evidência na RS prévia, e se a atualização da RS poderia melhorar a credibilidade dos achados¹³.

Para averiguar a necessidade de uma nova RS ou a atualização de uma RS já publicada, os autores devem fazer uma breve busca na literatura por RS já realizadas ou em andamento. Para isso, pode-se recorrer à base de RS da Cochrane Library, denominada Cochrane Database of Systematic Reviews (CDSR) (<https://www.cochranelibrary.com/cdsr/about-cdsr>), do Centre for Reviews and Dissemination (CRD) (<https://www.crd.york.ac.uk/CRDWeb/>) e do International Prospective Register of Systematic Reviews (PROSPERO) (<https://www.crd.york.ac.uk/PROSPERO/>), que são especializados em RS. Essas bases contêm referências de revisões que estão em andamento (acesso ao protocolo registrado). Adicionalmente, é desejável buscar protocolos de RS em bases de dados on-line de acesso aberto, tais como: Open Science Framework (OSF) (<https://osf.io/>) e Figshare (<https://figshare.com/>).

2.2 | Elaboração e registro do protocolo

O primeiro passo para a realização de uma RS é desenvolver um protocolo bem detalhado. Assim como em qualquer estudo primário, esse protocolo deve ser elaborado a priori com o objetivo de definir, de forma clara e transparente, todo o processo e os métodos para sua realização, incluindo as análises que serão realizadas. A elaboração do protocolo minimiza o risco de erros sistemáticos ou vieses da RS, introduzidos por decisões influenciadas pelos achados, além de ser uma forma de garantir que o estudo seja conduzido de acordo com o plano inicial do protocolo, independentemente de os resultados obtidos serem positivos ou não². Para elaborar o protocolo, é importante amadurecer a pergunta de pesquisa, confirmar a viabilidade da RS por meio de etapas-piloto de busca e seleção, e assegurar que existem estudos primários que respondem a pergunta de pesquisa da RS.

O protocolo descreve as etapas realizadas na revisão, como justificativa, objetivos, definição dos critérios de elegibilidade, bases de dados a serem pesquisadas, definição da estratégia de busca, processo de triagem e seleção dos estudos, processo de extração de dados, avaliação da qualidade metodológica dos estudos incluídos, avaliação da qualidade da evidência, plano de análises estatísticas, síntese de resultados e análises de sensibilidade e subgrupos². A elaboração do protocolo deve seguir as recomendações de relato do *Preferred Reporting Items for*

*Systematic Reviews and Meta-Analyses Protocol (PRISMA-P)*¹⁴.

O registro deve ser prospectivo, ou seja, deve ser realizado antes do início do processo de seleção dos estudos e, principalmente, da extração de dados. Os autores devem realizar o registro após elaborar e amadurecer o protocolo da revisão, para prever métodos factíveis e evitar alterações e/ou correções após o registro. Alterações posteriores devem ser identificadas e justificadas pelos autores na publicação final^{2,5}. O registro prospectivo do protocolo assegura transparência ao processo da RS, além de ajudar a identificar possíveis vieses de relato, já que, por meio do registro, é possível identificar diferenças entre os métodos e os desfechos pré-especificados no registro e os relatados na publicação¹⁵. Cerca de um terço das RS publicadas apresenta discrepância entre os desfechos relatados no registro da RS e os desfechos relatados na publicação^{16,17}, verificação que só foi possível pela disponibilidade do protocolo. O registro da RS evita a duplicação de RS por permitir a identificação de projetos em andamento. Isso evita duplicação de esforços e gastos desnecessários com a realização de revisão repetida¹⁵.

O registro do protocolo da RS também contribui para identificação do viés de publicação, já que é possível identificar revisões realizadas sobre determinado assunto, independentemente de elas estarem publicadas ou não¹⁵. Importante destacar que em torno de 50% dos protocolos registrados no PROSPERO não tiveram seus resultados publicados em revistas indexadas¹⁸.



O primeiro passo para a realização de uma RS é desenvolver um protocolo bem detalhado.

A plataforma mais utilizada de registro de revisões sistemática é o PROSPERO (<https://www.crd.york.ac.uk/prospero/>), o qual está ativo desde 2011. Instituições dedicadas à elaboração de RS, como Joanna Briggs Institute (https://jbi.global/systematic-review-register?combine=&items_per_page=10&page=10), Campbell Collaboration (<https://www.campbellcollaboration.org/>) e Cochrane (<https://www.cochrane.org/>), também registram os protocolos das RS elaboradas no âmbito dessas instituições. O protocolo de revisão pode ser publicado, ainda, em revistas científicas de acesso aberto (*open access*), o que permite a revisão, por pares, de métodos planejados, antes de serem concluídos ou em bases de dados *preprint* (ex.: OSF Registries - <https://osf.io/registries>; Figshare - <https://figshare.com/>).



2.3 | Ferramentas importantes para elaboração da revisão sistemática

2.3.1 | Ferramenta de relato

Um relato adequado e transparente de como foi realizada a RS contribui para melhor entendimento dos clínicos para a tomada de decisão correta, assim como os responsáveis pela inclusão ou não das tecnologias no sistema público de saúde. Informações importantes costumam não ser relatadas, ou ser relatadas de forma errônea, o que diminui a aplicabilidade dos achados da RS^{19, 20}.

O relato da RS deve seguir as recomendações do guia de redação *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*, o qual possui 27 itens para um relato completo e transparente, desde o título e resumo até a conclusão, além de sugestão de modelo de fluxograma do processo de seleção dos estudos (**Figura 6**)⁵. O checklist PRISMA, na versão em português, está disponível, no **Anexo A** desta diretriz (<http://www.prisma-statement.org/Translations/Translations>). As recomendações desse guia têm como objetivo melhorar a qualidade do relato das RS. Importante destacar que o checklist PRISMA não deve ser utilizado para avaliar a qualidade metodológica de RS.

Uma atualização do PRISMA está em fase final e deve ser publicada em breve. Portanto os autores de RS deverão utilizar essa versão mais atual assim que estiver disponível. Uma versão *preprint* pode ser encontrada no link <https://osf.io/preprints/metaarxiv/v7gm2/>.

2.3.2 | Ferramentas para avaliação da qualidade metodológica e risco de viés de RS

Existem ferramentas de avaliação de RS, e as mais utilizadas são a Assessing the Methodological Quality of Systematic Reviews (AMSTAR-2), para qualidade metodológica, e a Risk of bias Bias in Systematic Reviews (ROBIS)^{21, 22}, para risco de viés. A qualidade metodológica está relacionada ao rigor com que o estudo foi desenhado, conduzido, analisado e relatado para evitar erros sistemáticos, enquanto o risco de viés avalia quais falhas no desenho, na condução e na análise podem afetar os resultados de um RS²³. Apesar dessas ferramentas terem sido criadas para leitores avaliarem a qualidade metodológica ou o risco de viés das RS publicadas, elas também podem ajudar os autores a realizarem uma RS de melhor qualidade. Os autores devem conferir os itens e os processos avaliados por essas ferramentas e contemplá-los na RS.



3

CONDUÇÃO E RELATO DA REVISÃO SISTEMÁTICA



3.1 | Título

O título deve ser identificado como revisão sistemática (com ou sem meta-análise).

3.2 | Resumo estruturado

O resumo deve incluir: introdução, objetivos, fontes de dados, critérios de elegibilidade, participantes e intervenções, avaliação do risco de viés dos estudos incluídos e síntese de métodos, resultados, limitações, conclusões, achados principais e número de registro do protocolo da RS.

3.3 | Introdução

A introdução da RS deve fornecer contextualização do tema estudado, incluindo todos os fatores relevantes para a pergunta de pesquisa, justificando, portanto, sua elaboração⁷. É importante descrever a condição de saúde e a intervenção de interesse, incluindo uma síntese das evidências existentes. Caso exista alguma RS publicada anteriormente, esta deve ser referenciada na introdução, e os autores devem especificar como a proposta de uma nova revisão contribuirá para aumentar o conhecimento e preencher as lacunas identificadas na literatura científica⁴.

3.4 | Objetivos da revisão sistemática

A introdução deve ser finalizada com um objetivo abrangente de revisão que capte e se alinhe aos elementos do PICO. O objetivo deve ser claro em relação ao que a proposta da RS deseja alcançar⁴.

3.5 | Métodos

3.5.1 | Protocolo e registro

Os autores devem identificar se o protocolo da RS foi publicado e/ou registrado. Se este estiver disponível, forneça as informações relevantes e o número de registro.

3.5.2 | Definição da questão de pesquisa

Em uma RS de ECR, a questão de pesquisa a ser investigada deve ser clara e objetiva. A fim de orientar a formulação da(s) questão(ões) de pesquisa, convencionou-se estruturá-la(s) segundo os componentes do acrônimo PICO (**Quadro 1**), no qual cada letra representa um componente da questão, de acordo com os seguintes conceitos:

P – População: especifica qual é a população considerada pela RS.

I – Intervenção: define qual é a intervenção investigada.

C – Comparador: define o comparador ou o controle, que pode ser uma intervenção ativa, placebo, *sham*, nenhuma intervenção, entre outros.

O – Desfecho: proveniente da palavra em inglês *outcome*, define qual(is) será(ão) o(s) desfecho(s) investigado(s). É importante indicar os desfechos primários e os desfechos secundários.

Caso haja qualquer alteração da pergunta da RS durante a sua execução, essa alteração precisa estar explicitada e justificada na revisão.

Quadro 1 – Exemplos de perguntas de pesquisa para RS

1) Intervenção medicamentosa

Dexametasona é eficaz em pacientes com síndrome respiratória aguda grave ocasionada por covid-19 quando comparado ao cuidado usual?

P: pacientes com síndrome respiratória aguda grave ocasionada por covid-19.

I: dexametasona + cuidado de suporte.

C: cuidado de suporte.

O: desfecho primário: mortalidade, eventos adversos graves;

desfecho secundário: redução no tempo de ventilação mecânica.

(continua)

(conclusão)**2) Intervenção cirúrgica**

A cirurgia de mastectomia é eficaz e segura no tratamento da neoplasia mamária?

- P:** mulheres com neoplasia mamária.
I: cirurgia de mastectomia.
C: quimioterapia associada a radioterapia.
O: desfecho primário: sobrevida e eventos adversos graves;
desfecho secundário: taxa de resposta, recidiva local e eventos adversos não graves.

3) Intervenção fisioterapêutica

Exercícios do assoalho pélvico são eficazes no tratamento da incontinência urinária em mulheres idosas?

- P:** mulheres idosas com incontinência urinária de esforço.
I: exercícios do assoalho pélvico.
C: eletroestimulação transvaginal.
O: desfecho primário: melhora da incontinência urinária, da qualidade de vida e de eventos adversos graves;
desfecho secundário: aumento da força dos músculos do assoalho pélvico.

4) Intervenção educacional

Avaliação da eficácia de um programa de intervenção educacional no tratamento da depressão em mulheres no período perinatal.

- P:** mulheres no período perinatal com depressão.
I: programa de intervenção educacional (ferramentas que serão utilizadas para o fim devem estar descritas no protocolo) + prática clínica convencional.
C: prática clínica convencional.
O: desfecho primário: melhora da qualidade de vida, da gravidade da depressão e de eventos adversos graves;
desfecho secundário: níveis de ansiedade, níveis de estresse e eventos adversos não graves.

Fonte: Elaboração própria.

3.5.3 | Critérios de elegibilidade

Para uma RS de ECR, os critérios de elegibilidade devem estar baseados nos elementos do PICO e devem ser definidos a priori e registrados no protocolo da RS.

As características da **população** elegível devem ser identificadas levando-se em consideração a condição de saúde estudada. Podem ser definidas algumas informações, como idade, sexo, raça, etnia (se pertinente), e estágio ou gravidade da doença. O interesse pode ser focado em locais, contextos ou grupos, como hospital, instituto, área rural ou urbana, comunidade, profissionais, setor de saúde, entre outros. Os critérios de exclusão também devem estar explícitos nesta seção, sendo relacionados, por exemplo, com alguma condição de saúde específica ou severidade da doença^{4, 24}. Em situações nas quais a população de interesse da RS é uma das subpopulações consideradas no estudo primário, recomendamos que os autores da RS solicitem os dados separados para a subpopulação de interesse.

A **intervenção** deve ser identificada, assim como suas variações, em relação à dose/intensidade, à via de administração, à frequência e à duração²⁴.

Os **comparadores** elegíveis podem ser: outra intervenção ativa, a qual deve ser identificada de acordo com suas características e variações (dosagem/intensidade, via de administração, frequência, duração), placebo, *sham*, cuidados usuais, lista de espera, ou nenhum tratamento²⁴.

Os autores devem identificar se cointervenções serão permitidas, como o uso de analgésicos, corticosteroides, terapias hormonais, anti-inflamatórios, fisioterapia, cirurgias, treinamentos específicos para os pacientes, entre outros. Essas cointervenções podem influenciar no resultado do desfecho avaliado²⁴.

A especificação dos **desfechos** deve ser diferenciada em primários e secundários e deve prever suas possíveis mensurações (ex.: escala visual analógica, questionários, exames laboratoriais, entre outros), de preferência com ferramentas validadas. Podem ser incluídos os seguintes tipos de desfechos: sobrevivência e mortalidade, eventos clínicos (ex.: infarto do miocárdio), comportamentais (ex.: mudanças na dieta), relatados pelo paciente (ex.: qualidade de vida), eventos adversos. Os autores devem priorizar os desfechos de maior relevância clínica para a população. Marcadores bioquímicos ou fisiológicos são desfechos substitutos e devem ser ponderados

quanto à sua relação com desfechos importantes para os pacientes. No caso de desfechos com medidas repetidas, os autores necessitam especificar os tempos de seguimento de interesse, com as devidas justificativas.

Recomenda-se que sejam incluídos ao menos um desfecho primário de eficácia e um de segurança (com o número máximo de três desfechos primários). Os desfechos podem ser especificados com base na opinião de um especialista ou da literatura científica, por meio das recomendações de iniciativas voltadas para a identificação de desfechos relevantes, como Core Outcome Measures in Effectiveness Trials (COMET) (<http://www.comet-initiative.org/>) e Patient-Centered Outcomes Research (PCOR) (<https://www.ahrq.gov/pcor/index.html>).

3.5.4 | Fontes de informação e estratégias de busca

A busca por estudos deve ser abrangente, sensível e sistematizada, visando encontrar todas as evidências disponíveis que preencham os critérios de elegibilidade, com o objetivo de reduzir a possibilidade de viés de publicação. Os autores devem apresentar detalhadamente a estratégia de busca que foi elaborada para cada base de dados, identificando todos os termos utilizados, a data de busca e o número de estudos encontrados, para que essa busca seja reprodutível e o processo seja transparente.



A especificação dos desfechos deve ser diferenciada em primários e secundários.

3.5.4.1 Bases de dados

É recomendado que os autores realizem buscas em pelo menos três bases de dados gerais, sendo as principais: MEDLINE, CENTRAL, Embase e Lilasc. Além dessas bases, é importante que os autores pesquisem em bases de dados especializadas na área de interesse (quando relevante), por exemplo: AMED (terapias alternativas), CINAHL (enfermagem, biomedicina e ciências da saúde), PEDro (fisioterapia) e PsycINFO (psicologia e psiquiatria). No **Anexo B**, pode ser encontrada uma lista com as principais bases de dados para busca da evidência científica.

3.5.4.2 Literatura cinzenta

A literatura cinzenta é composta por quaisquer documentos técnicos ou científicos que não foram disponibilizados em bases de dados formais, incluindo, por exemplo, as teses e dissertações, os relatórios governamentais e os resumos de congressos que não foram publicados em suplementos de periódicos, ou seja, que possuem circulação restrita. Esses documentos podem conter informações importantes sobre resultados nulos ou negativos de estudos não publicados em periódicos científicos e devem ser contemplados na busca da RS, a fim de reduzir o risco de viés de publicação²⁵.

Para a busca de estudos da literatura cinzenta, os autores da RS podem consultar as seguintes bases:

- OpenGrey: <http://www.opengrey.eu/>
- MedRxiv: <https://www.medrxiv.org/>
- BiorXiv: <https://www.biorxiv.org/>
- Who Library Database: <http://kohahq.searo.who.int/>
- ProQuest: <https://www.proquest.com/libraries/academic/databases/>

- Catálogo de teses e dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes): <http://catalogo-deteses.capes.gov.br/catalogo-teses/#!/>
- Biblioteca Digital Brasileira de Teses e Dissertações: <https://bdtd.ibict.br/vufind/>

3.5.4.3 Bases de registros de ensaios clínicos

As bases de registros de ensaios clínicos são uma importante fonte de informação e podem fornecer resultados de potenciais estudos de interesse que não foram publicados em periódicos científicos. A busca nessas bases também funciona como um guia de estudos em andamento, para direcionar o momento certo para a realização de uma atualização da RS, caso haja estudos robustos a serem finalizados.

As mais relevantes bases de registros de ensaios clínicos são: clinicaltrials.gov, International Clinical Trials Registry Platform (ICTRP), The European Union Clinical Trials Register, Registro Brasileiro de Ensaios Clínicos (ReBEC).

3.5.4.4 Construção da estratégia de busca

A construção da estratégia de busca pode ser revisada por especialistas ou bibliotecários com experiência em realizar buscas em RS. Idealmente, as estratégias de buscas não devem ter restrições em relação ao idioma e ao ano de publicação. No caso de essas restrições serem utilizadas, os autores devem justificar essa opção no texto da RS. O primeiro passo para a construção da estratégia de busca é observar a pergunta de pesquisa referida pelo conceito PICO da RS, identificando as palavras-chave para a obtenção de estudos relevantes. No **Anexo C**, encontra-se um passo a passo de como construir uma estratégia de busca nas principais bases de dados.

É importante realizar uma estratégia-piloto em uma das bases de dados principais, geralmente a MEDLINE via PubMed, a fim de verificar se ela contempla os estudos principais que respondem a pergunta de pesquisa. A estratégia-piloto pode ser refinada e, uma vez que esteja adequada, pode ser utilizada como base para construção das estratégias de buscas nas demais bases de dados. Se necessário, os autores podem consultar o *Peer Review of Electronic Search Strategies – PRESS 2015 Guideline Evidence Based Checklist*, para auxiliar e/ou revisar a construção da estratégia de busca. Este é composto de perguntas e um guia, que podem direcionar os autores com relação ao uso de termos do conceito PICO na estratégia de busca, uso dos operadores booleanos, vocabulário controlado e não controlado e o uso de limites e filtros²⁶.

3.5.4.5 Busca sensível versus específica

A estratégia de busca é considerada sensível quando esta encontra a maioria dos estudos. Para a construção de uma estratégia de busca sensível nas bases de dados, devem-se utilizar tipicamente os termos referentes à população e à intervenção de interesse. A estratégia de busca torna-se específica quando não obtém muitos estudos irrelevantes para a RS²⁷. Utilizar termos referentes ao delineamento de estudo elegível na RS pode tornar a busca um pouco mais específica e, para essa finalidade, devem ser utilizados apenas os filtros validados e específicos para cada base de dados. Filtros validados para ECR podem ser acessados pelo website <https://work.cochrane.org/rct-filters-different-databases> (**Quadro 2**). No **Anexo D**, encontram-se exemplos de estratégias de buscas nas principais bases, com exemplos de filtros validados utilizados e número de referências obtidas com a utilização de cada filtro.

Quadro 2 – Exemplos de filtros para ECR

MEDLINE via PubMed (sensível e específica)	(randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR clinical trials as topic[mesh:noexp] OR randomly[tiab] OR trial[ti] NOT (animals[mh] NOT humans [mh]))
MEDLINE via PubMed (sensível)	(randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab] NOT (animals [mh] NOT humans [mh]))
EMBASE	‘crossover procedure’:de OR ‘double-blind procedure’:de OR ‘randomized controlled trial’:de OR ‘single-blind procedure’:de OR (random* OR factorial* OR crossover* OR cross NEXT/1 over* OR placebo* OR doubl* NEAR/1 blind* OR singl* NEAR/1 blind* OR assign* OR allocat* OR volunteer*):de,ab,ti
EMBASE (sensível)	random*:ab,ti OR (clinical NEXT/1 trial*):de,ab,ti OR ‘health care quality’/exp
EMBASE (específica)	(double NEXT/1 blind*):de,ab,ti OR placebo*:ab,ti OR blind*:ab,ti
EMBASE (sensível e específica)	random*:ab,ti OR placebo*:de,ab,ti OR (double NEXT/1 blind*):ab,ti

Fonte: Filtros sugeridos pela Cochrane <http://work.cochrane.org/rct-filters-different-databases>.

Os termos referentes ao comparador e/ou aos desfechos não devem ser utilizados nas estratégias de busca, pois tendem a ser ignorados no título e no resumo das publicações, e geralmente não são indexados como termos de vocabulários controlados. Porém, se os autores entenderem que essa abordagem deve ser utilizada, isso precisa ser justificado no texto da RS.

Os autores devem identificar se haverá restrição em relação ao ano de publicação ou ao idioma dos estudos. Não é recomendado restringir pelo ano de publicação sem que haja uma justificativa muito plausível (ex.: doenças recém-descobertas), já que uma busca restrita aos últimos cinco e dez anos podem identificar, respectivamente, apenas em torno de 25% e 46% dos estudos de interesse²⁸. Caso os autores optem por utilizar alguma restrição, esta deve ser devidamente justificada no texto da RS.

3.5.4.6 Vocabulário controlado e não controlado

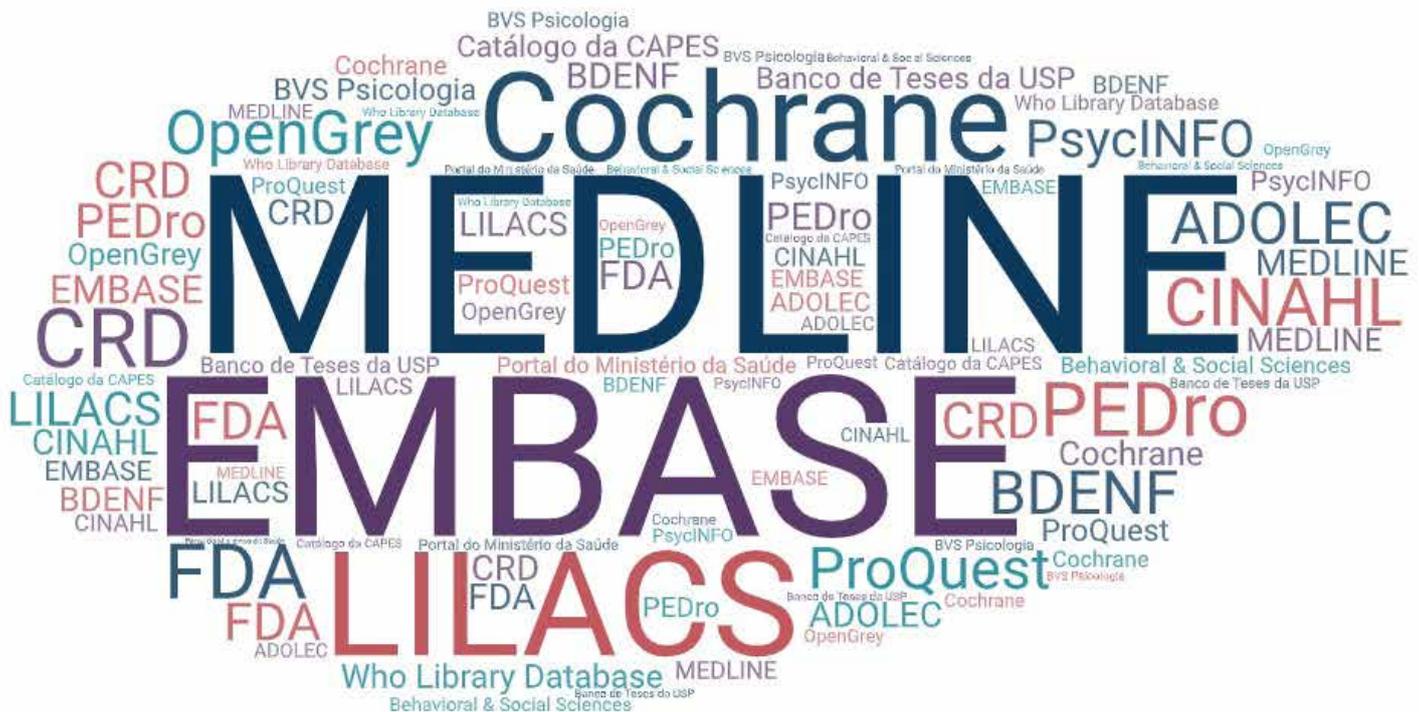
A estratégia de busca deve ser formulada utilizando-se termos de vocabulário controlado e termos de vocabulário não controlado²⁷. O vocabulário controlado é baseado em termos padronizados e descritores, que são atribuídos a determinado assunto e utilizados para indexar um estudo na base de dados. No quadro a seguir (**Quadro 3**), estão identificados os vocabulários controlados utilizados nas principais bases de dados.

Quadro 3 – Vocabulários controlados das principais bases de dados

3. Medical Subject Headings (MeSH)	https://www.ncbi.nlm.nih.gov/mesh/ Cochrane Library https://www.cochranelibrary.com/advanced-search/mesh
Emtree	Embase https://www.embase.com/#emtreeSearch/default
Descritores em Ciências da Saúde (DeCS)	Lilacs e BVS http://decs.bvs.br/

Fonte: Elaboração própria.

É importante identificar, na estratégia de busca, o termo referente ao vocabulário controlado utilizado, por exemplo: **Cochrane Library**: MeSH descriptor: [Celiac Disease] explode all trees; **MEDLINE via PubMed**: (“Celiac Disease”[Mesh]); **Lilacs via Biblioteca Virtual em Saúde**: ((mh:(Doença celíaca)).

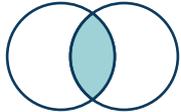
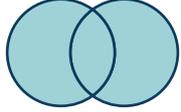
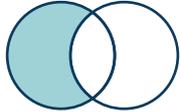


O vocabulário não controlado é baseado no uso de termos livres que usualmente estão disponíveis no título ou no resumo das publicações. Os autores devem buscar por sinônimos, acrônimos, ou termos relacionados à população e à intervenção de interesse, no sentido de tornar a busca mais abrangente. Cada base de dados fornece sinônimos do vocabulário controlado e alguns termos relacionados. No MEDLINE, por exemplo, para cada termo MeSH, há uma lista de termos relacionados, chamada de *entry terms*, os quais devem ser incorporados à estratégia de busca. Para melhor identificação dos sinônimos, recomenda-se a utilização de dicionários médicos e a consulta ao sítio do Thesaurus (www.thesaurus.com). Na base de dados Lilacs, é importante utilizar os sinônimos do termo descritor de assunto em todos os idiomas disponíveis (espanhol, inglês e português) no intuito de deixar a busca mais sensível.

3.5.4.7 Operadores booleanos

Após definir os termos que serão utilizados na estratégia de busca, os autores devem recorrer aos operadores booleanos para a combinação deles. Os operadores booleanos são universais e servem para realizar as combinações demonstradas no **Quadro 4**.

Quadro 4 - Operadores booleanos

<p style="text-align: center;">AND (intersecção)</p>	<p>Relação entre dois ou mais assuntos.</p> <p>Ex.: tromboembolismo AND varfarina.</p> <p>Recupera estudos que contenham tromboembolismo e varfarina.</p>	
<p style="text-align: center;">OR (união)</p>	<p>Relação entre termos do mesmo assunto.</p> <p>Ex.: <i>celiac disease</i> OR <i>celiac sprue</i>.</p> <p>Recupera estudos que contenham <i>celiac disease</i> ou <i>celiac sprue</i>.</p>	
<p style="text-align: center;">NOT (exclusão)</p>	<p>Demonstra que o primeiro assunto deve ser incluído, e o segundo excluído.</p> <p>Ex.: <i>diabetes</i> NOT <i>animals</i>.</p> <p>Recupera estudos de diabetes que não tenham sido realizados em animais.</p>	

Fonte: Elaboração própria.

Os termos utilizados na busca referentes a cada conceito do PICO (por exemplo, termos relacionado à população) devem ser combinados com o uso do operador booleano “OR”, para assegurar a sensibilidade dentro de cada conceito. Os resultados para cada conceito são então combinados usando o operador booleano “AND” no resultado da busca final; por exemplo, na combinação dos termos da população com os termos da intervenção²⁷. O uso do operador “NOT” normalmente deve ser evitado no contexto de uma RS.

3.5.4.8 Busca manual de estudos

Os autores devem analisar as referências citadas nos estudos recuperados, em revisões sistemáticas relevantes identificadas e em anais de congresso importante da área, para verificar se há algum estudo relevante que não foi contemplado pela estratégia de busca previamente realizada⁴. Quando um estudo é encontrado, os autores devem incluí-lo de forma manual para a realização do processo de extração de dados, lembrando que ele deve estar identificado no fluxograma de seleção dos estudos. É recomendado, também, que os autores da RS realizem contato com especialistas da área para checar potenciais estudos relevantes para a RS.

3.5.4.9 Gerenciadores de referências

Após a finalização das estratégias de busca em cada base de dados, é possível exportar todos os resultados dos estudos obtidos. Nesse sentido, sugere-se exportar no formato “RIS format”, pois é o mais aceito nos gerenciadores de referência.

Os gerenciadores de referência facilitam a organização das referências e a unificação delas, além de serem ferramentas importantes para a exclusão de estudos duplicados advindos de várias bases de dados pelo processo de busca. Os autores da RS podem utilizar ferramentas gratuitas, como: Mendeley, Zotero, EndNote web, entre outras; ou podem optar pelo uso de ferramentas comerciais, como: EndNote ou Reference Manager. As plataformas de seleção de estudos (detalhadas adiante) também possuem a finalidade de identificação de duplicatas.

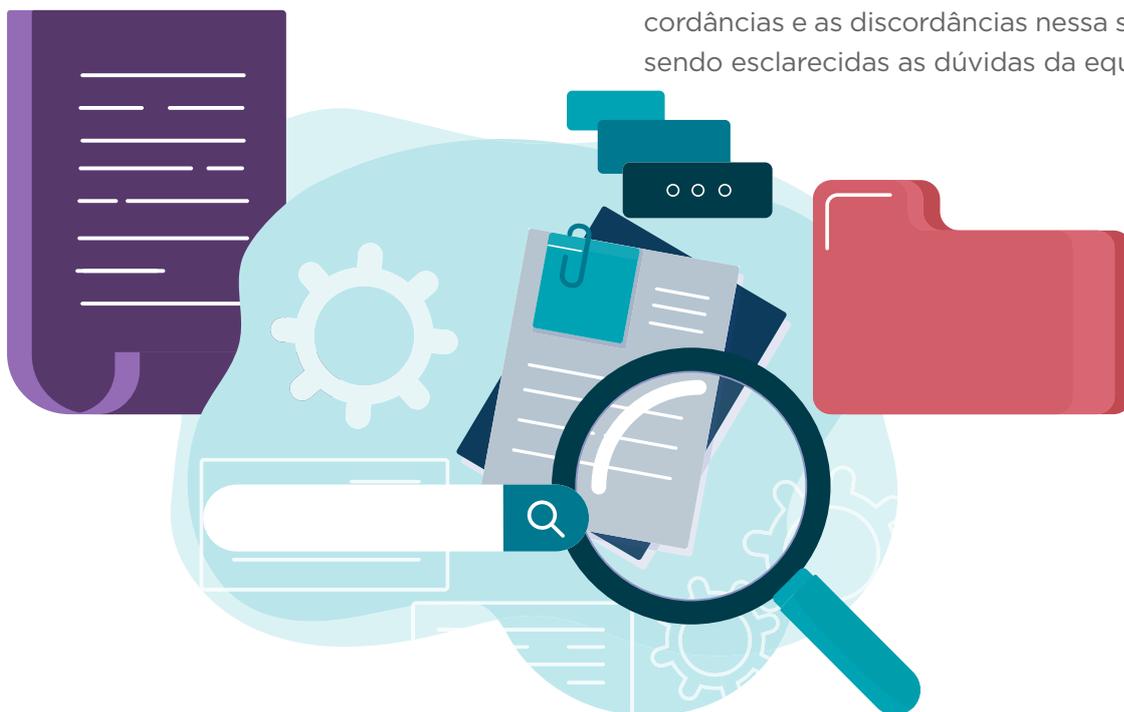
3.5.5 | Seleção dos estudos

A decisão pela inclusão ou não dos estudos deve ser baseada nos critérios de elegibilidade da RS. O processo de seleção (triagem de títulos e resumos; leitura de textos completos) deve ser realizado por pelo menos dois revisores, de modo independente, e, em caso de discordância, deve ser consultado o parecer de um terceiro avaliador. Esse formato reduz o risco de erros e a influência de viés de uma única pessoa, e deve ser utilizado tanto na fase de triagem de títulos e resumos quanto na fase de leitura de textos completos.

Algumas ferramentas podem auxiliar no processo de seleção dos estudos, como a plataforma gratuita Rayyan, ou ferramentas comerciais, como o Microsoft Excel, Covidence, EPPI-Reviewer, entre outras⁴. Por meio do website: <http://systematicreviewtools.com/>, é possível encontrar uma lista de ferramentas que auxiliam no processo de seleção dos estudos.

3.5.5.1 Triagem de títulos e resumos

A primeira fase da seleção dos estudos pode ser identificada como a triagem de títulos e resumos. Os autores da RS devem identificar, nessas seções, se o estudo atende aos critérios de inclusão da RS, observando se a população e a intervenção são de interesse, se os participantes não possuem nenhum tipo de condição de saúde que deveria ser excluída, se o delineamento do estudo é elegível, entre outros aspectos. No caso de os autores ficarem em dúvida se o estudo é elegível, pela falta de informações provenientes no título ou no resumo, devem considerar o estudo elegível para a segunda fase de seleção (leitura de textos completos). Recomenda-se realizar calibração da seleção, com os revisores, com um número reduzido de publicações, para confirmar a compreensão dos critérios de elegibilidade. Essa calibração pode ocorrer por meio de reunião em que são verificadas as concordâncias e as discordâncias nessa seleção inicial, sendo esclarecidas as dúvidas da equipe.



3.5.5.2 Leitura de textos completos

Os estudos selecionados como elegíveis na primeira fase de seleção (triagem de títulos e resumo) devem ser analisados detalhadamente e com cautela. Nesse intuito, os autores devem fazer uma busca para obtenção do texto completo e outras potenciais fontes de informação do mesmo estudo (ex.: registro de ensaios clínicos e outras publicações), a fim de identificar sua elegibilidade. Da mesma forma que foi realizada na primeira fase de seleção, os autores devem buscar, no texto, palavras-chave que identifiquem os critérios de elegibilidade com relação ao conceito PICO da RS e ao tipo de delineamento de estudo. É importante observar se os resultados dos desfechos de interesse foram apresentados, de modo a verificar a elegibilidade. Em algumas situações, os autores podem optar por não considerar desfechos como critérios de elegibilidade; nesse caso, eles não seriam uma razão para exclusão após leitura de texto completo.

Na fase de leitura de textos completos, os autores devem apontar a justificativa de exclusão dos estudos. Essas justificativas em geral podem ser identificadas no fluxograma de seleção como: outra condição de saúde, outro delineamento de estudo, outra intervenção, entre outros critérios de exclusão relevantes. No entanto é mandatório que haja uma lista com os estudos excluídos e a identificação da razão para exclusão de cada um deles.

Ao final do processo de seleção, os autores deverão ter identificado os estudos excluídos, os estudos incluídos e os estudos em andamento.

3.5.6 | Extração de dados

Após o processo de seleção dos estudos, deve-se realizar a busca de todos os documentos adicionais referentes aos estudos incluídos, ao registro, aos resumos publicados, aos protocolos, às erratas, às cartas, ao material suplementar e às informações de agência reguladoras (ex.: US Food and Drug Administration [FDA]), se houver.

3.5.6.1 Fatores que devem ser considerados antes da extração dos dados

Presença de erratas

É importante examinar erratas, pois podem conter informações importantes sobre o estudo, como estudos considerados fraudulentos, que podem culminar na exclusão do estudo na RS ou na meta-análise.

Múltiplas publicações do mesmo estudo

Refere-se à prática de fragmentar a apresentação de resultados de uma mesma pesquisa, reduzindo-os às suas mínimas partes, as quais serão apresentadas em distintas publicações como se fossem resultados de pesquisas independentes²⁹.

Publicações do mesmo estudo não devem ser consideradas como vários estudos. No entanto as publicações secundárias de um estudo não devem ser descartadas, pois podem conter informações importantes sobre o delineamento e a condução dele. A publicação principal normalmente é a que contém o maior número de informações sobre a metodologia e deve apresentar os resultados dos desfechos de interesse.

Contato com autores

Em caso de dúvidas, os autores dos estudos primários devem ser contatados, sendo que o contato pode ser encontrado por meio de e-mail contido na publicação, listas de funcionários de universidades ou instituições, diretórios de membros de sociedades profissionais, ResearchGate, LinkedIn, pesquisa geral na web ou contato com outros autores. Mais de uma tentativa de contato deve ser feita e todas elas devem ser registradas em uma planilha de controle contendo: nome dos autores, e-mail de contato dos autores, data de envio do e-mail, conteúdo do e-mail (dados solicitados), data do retorno dos autores e conteúdo das respostas dos autores.

3.5.6.2 Processo de extração

Número de avaliadores

A extração de dados deve ser feita por dois avaliadores independentes e, posteriormente, deve ser realizado o consenso. Em caso de discordância, é preciso solicitar o parecer de um terceiro avaliador^{5, 27}. A extração de dados por dois avaliadores independentes normalmente resulta em uma taxa de erro menor, comparada à extração realizada por um revisor e checada por um segundo³⁰.

Ferramentas de extração

A escolha da ferramenta de extração utilizada na RS dependerá de sua complexidade e do número de estudos que serão incluídos na revisão³¹. Há uma lista de softwares que podem ser utilizados para o processo de extração, na página da web: <http://systematicreviewtools.com/index.php>.

Um dos grupos da Cochrane Development, Psychosocial and Learning Problems (Cochrane) disponibilizou um formulário-padrão para extração de ensaios clínicos randomizados, o qual pode ser acessado pelo seguinte website: <https://dplp.cochrane.org/data-extraction-forms>. Esse formulário pode ser utilizado como base para a construção de um formulário específico para a pergunta de pesquisa. Outra maneira de criar formulários seria de forma eletrônica, com o Google Forms, por exemplo; ou por meio de planilhas de dados pelo Microsoft Excel, versão desktop ou on-line. O Microsoft Excel permite realizar consensos automatizados com a utilização de fórmulas ligando as informa-

ções das planilhas dos dois pesquisadores responsáveis pelo processo de extração. Existem sistemas de dados personalizados ou comerciais (por exemplo, Covidence, EPPI-Reviewer, Systematic Review Data Repository – SRDR) que permitem a criação de formulários on-line, a entrada de dados por vários usuários, o compartilhamento de dados e o gerenciamento eficiente destes, porém podem ser mais custosos.

Formulário para extração de dados

Os autores devem criar um formulário de extração com questões relevantes (**Quadro 5**). Se necessário, pode ser consultado um especialista na área. Os dados importantes de se extrair são preestabelecidos por meio do formulário pré-planejado e são referentes às características clínicas e metodológicas dos estudos e à avaliação dos desfechos de interesse da revisão sistemática.

Sugere-se realizar perguntas fechadas, com alternativas de escolha, antecipando e estruturando as possíveis respostas, e incluí-las em uma categoria. Deve-se evitar pedir aos extratores de dados para resumi-los em texto e fazer perguntas de forma que a resposta possa ser deixada em branco. Para evitar a resposta em branco, poderiam ser incluídas as opções: “não aplicável”, “não relatado” e “não é possível dizer”.

Para as questões que envolvam variáveis com possibilidade de diferentes unidades de medida – como exemplo, a duração do estudo –, deve haver uma padronização na extração para todos os estudos, realizando conversões, quando necessário.

Quadro 5 – Exemplo de questões que podem ser incluídas no formulário de extração

Identificação do estudo
ID do estudo Primeiro autor Ano de publicação Dados bibliográficos (se pertinente) País onde o estudo foi realizado Número do registro do protocolo do estudo Nome das intervenções Tamanho da amostra Duração do estudo Critérios de elegibilidade
Financiamento
O estudo recebeu suporte financeiro de empresa comercial? A empresa comercial está envolvida na execução do estudo? Foi declarada a existência de conflitos de interesse?
Desenho do estudo
Qual o desenho do estudo? Quantos braços tem o estudo? Como foi realizada a randomização? Qual método de sigilo de alocação foi utilizado? Houve cegamento em quais níveis?
Análises do estudo
A análise foi realizada por intenção de tratar (ITT – do inglês, <i>intention-to-treat</i>), ITT modificada, por protocolo ou outra? Foi realizada a imputação de dados? Se sim, qual o método de imputação
Características da linha de base
Número de pacientes relatado na linha de base Tempo de diagnóstico ou duração dos sintomas Número de mulheres/homens Etnia (se pertinente) Média de idade Média do índice de massa corporal (se pertinente)

(continua)

(conclusão)

Graduação/estadiamento da doença (se houver)
Número de pacientes com Grau 1 Número de pacientes com Grau 2
Tratamentos concomitantes
Foram permitidos tratamentos concomitantes? Quais?
Instrumento de avaliação do desfecho de interesse
Qual o instrumento utilizado? No caso de escalas, os valores mais elevados indicam piora do sintoma? No caso de escalas, qual o valor mínimo? No caso de escalas, qual o valor máximo?
Nome da intervenção (Repetir para cada grupo de intervenção)
Tipo de intervenção Número de indivíduos randomizados Dosagem/frequência/número de aplicações Duração máxima de tratamento Duração máxima de seguimento

Desfechos Contínuos

Informações em nível de cada período de avaliação em cada braço
Linha de base/tempo de seguimento Número de pacientes Média Desvio-Padrão Outras medidas de precisão e dispersão (se necessário) Foram utilizadas aproximações de gráficos?

Desfechos Categóricos

Número total de pacientes
Número total de pacientes que experimentaram o desfecho Número de pacientes que não experimentaram o desfecho

Fonte: Elaboração própria.

Após os autores realizarem a construção do formulário de extração, sugere-se a criação de um *codebook* (guia codificado) para as questões que possam gerar dúvidas durante a extração, a fim de que o consenso seja facilitado e os extratores tenham a mesma interpretação sobre a pergunta. Um exemplo é identificar as perguntas da extração com códigos ou números, explicar a questão e dar exemplos de respostas, como o exemplo do **Quadro 6**. Se os autores forem utilizar o Microsoft Excel para realizar a extração de dados, a explicação da pergunta e o exemplo podem ser inseridos como um comentário (inserir → novo comentário) na célula da pergunta, para facilitar a compreensão dos extratores.

Após a criação do *codebook* e do formulário de extração, é importante realizar um treinamento com os revisores e uma extração-piloto de alguns estudos, de modo a verificar se não é necessário modificar ou acrescentar alguma questão relevante.

Quadro 6 - Exemplo de *codebook* (guia codificado)

Questão	Número da questão	Código da pergunta	Explicação	Exemplo
Nome dos braços	Questão 10	nome_bra	Colocar os braços na seguinte ordem: 1. Placebo, <i>sham</i> , lista de espera, sem intervenção 2. Intervenção intravenosa 3. Intervenção intramuscular 4. Intervenção subcutânea 5. Intervenção oral 6. Intervenção cutânea Colocar em ordem alfabética; da menor dosagem para maior dosagem; da menor frequência para maior frequência	1. Placebo 2. Cloridrato de tramadol 3. Eletroterapia

Fonte: Elaboração própria.

Processo de extração em pares

É recomendado que cada avaliador extraia os dados dos estudos de maneira independente e que não tenham acesso às respostas do outro avaliador até a fase do consenso. Quando o consenso for realizado, os autores devem manter suas respostas originais e criar um novo item com a resposta consenso. No caso de discordância das respostas, os autores devem discutir e entrar em um consenso, revisitando o estudo que foi utilizado para extração ou analisando materiais suplementares e arquivos de registros dos estudos. Caso ainda haja discordância, os autores devem solicitar o parecer de um terceiro avaliador. A calibração da extração com a equipe responsável é recomendada, comparando a extração dos primeiros estudos para aprimoramento do formulário e esclarecimento do processo.

Extração de dados de figuras

No processo de extração, os revisores podem se deparar com resultados de desfechos apresentados em figuras, como gráficos. Primeiramente, é importante verificar se os mesmos dados não foram apresentados em uma tabela no estudo principal ou nos outros arquivos do estudo como material suplementar, registro do estudo, resumos de congresso, outras publicações do mesmo estudo, entre outros. Sempre que possível, os dados devem ser extraídos de tabelas ou do texto do estudo, quando disponível. Se não houver os dados disponíveis, existem métodos de estimar os dados apresentados nas figuras dos estudos.

É importante verificar a padronização da escala métrica do gráfico e os valores mínimos do eixo x e y. Se o gráfico não apresentar escala padronizada, não será possível mensurar com exatidão os dados, e uma mensuração nessa figura poderia inferir resultados equivocados. Se a figura estiver com escala padronizada, mas iniciar de um valor que não seja o zero, ainda assim será possível mensurar o dado, porém é necessário atenção no momento de extração.

A extração de dados de figuras pode ser feita por meio de uma régua ou por meio de softwares, como o Digitizer, Digitizelt, WebPlot-Digitizer, Engauge, ycasd, Dexter, GetData Graph Digitizer³². O software é mais conveniente, mais rápido e tem maior acurácia do que o método visual ou a utilização de régua³³. De todo modo, o método usado deve ser detalhado no protocolo da RS e no texto final dela.

Tipos de dados

É importante identificar o tipo de dado que será extraído de acordo com o tipo de desfecho de interesse³⁴. A seguir, estão identificados os cinco tipos mais comuns:

- Dicotômicos (ou binários): cada desfecho individual apresenta somente duas respostas categóricas.
- Contínuos: cada desfecho individual é mensurado em quantidade numérica.
- Dados ordinais (incluindo mensuração de escalas): cada desfecho individual possui uma ordem ou é gerado pela soma de respostas categóricas.
- Contas e taxas: calculadas a partir da contagem do número de eventos experimentados por cada indivíduo (ex.: número de pacientes que apresentaram dor abdominal).
- Tempo para evento (tipicamente sobrevida): dados que analisam o tempo até a ocorrência do evento, mas nem todos os indivíduos do estudo o vivenciam (dados censurados).

Medidas de efeito

Estimativas de efeito descrevem a magnitude do efeito da intervenção, mostrando a diferença de efeito entre os grupos para determinado desfecho. Para medidas de efeito relativas (razão de risco e razão de chance), o valor 1 (um) representa que não há diferença entre os grupos. Para medidas de diferença, o valor 0 (zero) representa sem diferença entre os grupos. Valores superiores ou inferiores a esse valor “nulo” podem indicar efeito benéfico ou dano de uma intervenção experimental, dependendo de como as intervenções foram ordenadas³⁴.

Dados de desfechos dicotômicos

Os dados de desfechos dicotômicos (binários) surgem quando o desfecho de cada participante só tem duas possibilidades, por exemplo a mortalidade. As medidas de efeito para desfechos dicotômicos são:

- Razão de risco ou risco relativo (*risk ratio or relative risk* - RR).
- Razão de chance (*odds ratio* - OR).
- Diferença de risco (*risk difference or absolute risk reduction* - RD).
- Número necessário para tratar (*number needed to treat* - NNT).

Quadro 7 - Cálculo do risco relativo, razão de chance, diferença de risco e número necessário para tratar, com uma tabela de desfecho dicotômico

	Número de indivíduos com evento	Número de indivíduos sem evento	Número total de indivíduos
Intervenção experimental	A	B	N ₁
Intervenção comparadora	C	D	N ₂

Fonte: Elaboração própria.

Risco relativo = risco de evento no braço experimental / risco de evento no braço comparador = $(A/N_1) / (C/N_2)$

Razão de chance = chance de evento no braço experimental / chance de evento no braço comparador = AD / BC

Diferença de risco = risco de evento no braço experimental - risco de evento no braço comparador = $(A/N_1) - (C/N_2)$

Número necessário para tratar = $1 / \text{diferença de risco} = 1 / (A/N_1) - (C/N_2)$

Interpretação das medidas de efeito de dados dicotômicos

Risco relativo

O risco relativo (RR) corresponde à razão de risco de um evento nos dois grupos. O valor 1 indica que as estimativas de efeito são iguais para ambos os grupos. No caso do RR ser maior que 1 (por exemplo, 2), isso implica que eventos no braço com intervenção experimental são duas vezes mais prováveis que no braço com intervenção comparadora. Outra interpretação é que o risco de eventos aumenta em $100 \times (RR-1) \% = 100\%$ com o uso da intervenção experimental. Se o resultado corresponder a um número decimal menor que 1 (por exemplo, 0,25), isso indicaria que a probabilidade de um evento com o uso da intervenção experimental é um quarto da intervenção comparadora, ou ainda que a intervenção experimental reduziu o risco de evento em 75%.

Razão de chance

A razão de chances (do inglês, *odds ratio*) é definida como a razão entre a chance de um evento ocorrer no grupo experimental e a chance de um evento ocorrer no grupo comparador. Uma razão de chances 1 indica que o evento tem a mesma chance de ocorrer nos dois grupos. Uma razão de chances maior que 1 indica que o evento tem maior chance de ocorrer no grupo experimental e, em caso de ser menor que 1, a chance seria menor no grupo experimental³⁴.

Diferença de risco

A diferença de risco descreve a diferença no risco de eventos observados entre o grupo experimental e o grupo comparador. Para um indivíduo, seria interpretado como a diferença estimada na probabilidade de haver um evento³⁴.

Número necessário para tratar

O número necessário para tratar (NNT) é definido como o número esperado de pessoas que precisam receber a intervenção experimental em vez da comparadora, para que uma pessoa adicional incorra ou evite um evento (dependendo da direção do resultado)³⁴. Como exemplo, um NNT 10 pode ser interpretado da seguinte forma: é esperado que uma pessoa adicional incorra em um evento para cada dez pessoas que recebem a intervenção experimental em vez da intervenção comparadora.

Extração de dados de desfechos dicotômicos

É essencial extrair o número total de participantes avaliados para os desfechos dicotômicos de cada grupo, o número de participantes que experimentaram o evento de interesse e o nú-

mero de participantes que não experimentaram o evento de interesse descrito nos desfechos da RS. Esses números serão utilizados para a construção da tabela 2x2, a qual está demonstrada no **Quadro 7**.

Dados de desfechos contínuos

Medidas de efeito de desfechos contínuos

Três métodos de estatísticas sumárias podem ser realizados para meta-análise de desfechos contínuos: a diferença de médias (DM), a diferença de médias padronizada (DMP) e a razão de médias, sendo as duas primeiras as mais comumente utilizadas. Esses podem ser calculados a partir de dados de estudos individuais de mensurações pós-intervenção ou mensurações de mudança em relação à linha de base (do inglês, *change from baseline*). Outra forma de mensurar o efeito é por meio da razão de médias³⁴.

Diferença de médias

A DM mensura a diferença absoluta entre o valor da média nos dois grupos. Este método pode ser utilizado quando os estudos utilizarem a mesma escala para mensurar o desfecho de interesse³⁴.

Diferença de médias padronizada

A DMP é utilizada quando os estudos avaliam o mesmo desfecho por meio de escalas ou métodos diferentes. Nesse caso, é necessário padronizar os resultados por meio de uma escala uniforme antes de combinar os resultados. A DMP expressa o tamanho do efeito da intervenção em relação à variabilidade entre os participantes na mensuração do desfecho observado no estudo³⁴.

$$\text{Diferença de médias padronizada} = \frac{\text{diferença na média dos desfechos entre os grupos}}{\text{desvio-padrão do desfecho entre os participantes}}$$

Se as escalas analisadas apontarem para sentidos diferentes – por exemplo, uma escala aumenta com a gravidade da doença e a outra diminui –, é necessário multiplicar os valores médios de um conjunto de estudos por -1, para garantir que todas as escalas apontem para a mesma direção, antes da padronização. Esse ajuste deve ser descrito na seção de métodos da RS. O desvio-padrão não precisa ser modificado.

Razão de médias

A razão de médias estima qual o multiplicador do valor médio do grupo experimental em relação ao grupo comparador; por exemplo, uma razão de médias de 3 indica que a pontuação média dos participantes que receberam a intervenção é três vezes maior do que a do grupo comparador. Esta medida pode ser utilizada na meta-análise somente quando os resultados forem positivos – por exemplo, valores pós-intervenção –, e não para valores negativos, como valores de mudança da linha de base³⁴.

A razão de médias pode ser utilizada para comparar resultados de estudos que utilizaram escalas diferentes, desde que estas tenham limites inferiores comparáveis.

Extração de dados de desfechos contínuos

Para realizar a meta-análise de desfechos contínuos, é necessário extrair dos estudos a média, o desvio-padrão e o número total de participantes avaliados (N) em cada braço de intervenção para o desfecho de interesse. No entanto alguns estudos informam apenas valores de mediana, erro-padrão, intervalo de confiança (IC), intervalos interquartis, e mínimo e máximo.

Existem alguns cálculos matemáticos que podem auxiliar na transformação dos valores fornecidos pelos estudos individuais, os quais estão expostos a seguir, considerando um IC de 95%:

Desvio-Padrão

$$\text{Erro-Padrão} = \frac{(\text{limite superior do IC} - \text{limite inferior do IC})}{3,92}$$

$$\text{Desvio-Padrão} = \sqrt{N} \times \frac{(\text{limite superior do IC} - \text{limite inferior do IC})}{3,92}$$

No caso de os estudos apresentarem um IC de 90%, o valor 3,92 do cálculo matemático exposto acima deve ser substituído por 3,29. Já no caso de o IC ser 99%, esse valor deve ser substituído por 5,15.

Se N for menor que 100 participantes por grupo, então o valor 3,92 deve ser substituído por números específicos da distribuição t, o qual pode ser obtido por tabelas de distribuição t com graus de liberdade igual ao tamanho amostral menos 1.

Nas situações em que os ECR apresentam apenas dados de mediana, intervalo interquartil, ou valor mínimo e valor máximo, existe a possibilidade de cálculo de conversão deles para média e desvio-padrão³⁵. Para maiores informações sobre os cálculos matemáticos e como proceder no caso de dados ausentes nos estudos individuais, sugere-se consultar o Capítulo 6 do *Handbook da Cochrane*³⁴. Os autores da revisão sistemática não precisam se preocupar em realizar esses cálculos manualmente, pois existem algumas calculadoras disponíveis (como a do software gratuito Review Manager) que realizam esses cálculos/conversões automaticamente.

3.5.7 | Avaliação do risco de viés

Viés é definido como um erro sistemático, ou desvio da verdade, nos resultados³⁶. Autores das RS devem sempre investigar o quanto os ECR incluídos procuraram excluir ou minimizar vieses no seu delineamento, execução e análise dos dados que possam ter afetado a validade interna e a confiança dos seus achados.

A avaliação do risco de viés dos estudos primários incluídos em uma RS é importante para

ênfatar o risco desses vieses em superestimar (na maioria das vezes) ou subestimar o efeito real das intervenções, uma vez que as conclusões de uma RS dependem dos resultados dos estudos incluídos. Assim, se os resultados dos ECR estão enviesados, as sínteses da revisão podem levar a conclusões inadequadas e, conseqüentemente, a tomadas de decisões equivocadas. Apesar de haver evidência empírica de que determinados aspectos metodológicos levam à presença de vieses nos estudos, não é possível determinar a extensão da influência deles nos resultados do estudo³⁷.

Existem algumas ferramentas (por exemplo, Cochrane Risk of Bias – RoB, PEDro, Joanna Briggs – JBI Critical Appraisal Checklist for Randomized Controlled Trials) que têm sido utilizadas para avaliar a qualidade metodológica e/ou o risco de viés de ECR no contexto de RS. Porém algumas ferramentas são escalas ou checklists (ex.: Jadad Maastricht) que contêm itens que não são relacionados à validade interna (portanto não devem ser utilizadas) – por exemplo, se os estudos tinham poder estatístico ou não (item relacionado à precisão dos resultados, não à presença de viés) ou se os critérios de inclusão ou exclusão foram claramente descritos (item relacionado à aplicabilidade, não à validade)³⁸.

3.5.7.1 Ferramenta Cochrane Risk of Bias (RoB 2)

Esta atualização da Diretriz de Revisão Sistemática de Ensaios Clínicos Randomizados recomenda que a avaliação do risco de viés dos estudos incluídos seja realizada pela nova versão da ferramenta Cochrane Risk of Bias (RoB 2)²³. Essa ferramenta é composta por cinco domínios, os quais são baseados em evidência empírica, mas considerações teóricas também foram levadas em conta para a escolha deles.

Os domínios avaliados na ferramenta RoB 2 são: viés no processo de randomização; viés devido a desvios das intervenções pretendidas; viés devido a dados faltantes dos desfechos; viés na mensuração dos desfechos; e viés na seleção dos resultados relatados²³. Houve mudança dos nomes dos domínios em relação à versão anterior da ferramenta, mas os conceitos são parecidos.

Na nova versão da ferramenta, foram adicionadas questões sinalizadoras com opções de resposta (Sim; Provavelmente Sim; Provavelmente Não; Não; Nenhuma Informação) dentro de cada domínio, para apoiar o julgamento do risco de viés de cada domínio em “baixo risco de viés”, “algumas preocupações” e “alto risco de viés”. A ferramenta Rob 2 possui um algoritmo no Microsoft Excel (<https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2>) que usa as respostas das perguntas sinalizadoras para propor um julgamento de cada domínio e facilitar o processo de forma automatizada²³. Nesta diretriz, a ferramenta foi traduzida de forma livre para ajudar os autores de RS na sua utilização e encontra-se disponível no **Anexo A**.

Antes de utilizar a ferramenta RoB 2, os autores precisam definir se o objetivo da RS é avaliar o efeito da atribuição para intervenção no início do estudo, independentemente de as intervenções terem realmente sido recebidas no decorrer do estudo (efeito por intenção de tratar), ou o efeito da adesão à intervenção, conforme o protocolo (efeito por protocolo). O julgamento adequado do domínio, desvio das intervenções pretendidas, irá variar de acordo com o interesse da RS. Isso é importante quando os participantes nos ensaios clínicos não recebem a intervenção atribuída ou quando há desvios das intervenções planejadas, como baixa adesão ao tratamento. A maioria das RS deverá optar em avaliar o efeito da atribuição

para a intervenção, em vez do efeito da adesão à intervenção. Isso é devido a problemas com os métodos comumente usados até o momento (análise por protocolo simples e análise *as-treated*) em estimar de forma adequada o efeito da adesão à intervenção²³.

Domínios

Processo de randomização

O domínio contempla o método de randomização usado, o método de sigilo de alocação e, por fim, diferenças clinicamente relevantes entre características importantes da linha de base dos grupos de intervenção²³.

Sequência de randomização: métodos de randomização adequados são aqueles que incluem o elemento de chance no processo que determina a sequência com que os participantes serão alocados para os grupos de intervenção. A sequência de randomização deve ser conduzida por uma pessoa que não esteja diretamente envolvida com a avaliação e o tratamento dos participantes do estudo.

- Métodos considerados adequados: tabela com números aleatórios; geração de números por um programa de computador; sorteio utilizando moedas ou dados; e sorteio de cartões ou envelopes com reposição (cartões e envelopes retornam para o conjunto após cada sorteio).
- Métodos considerados inadequados (não possui componente aleatório): sequência gerada por data ímpar ou par; data de nascimento ou de admissão; número de registro hospitalar; sorteio sem reposição (números ou cartões são excluídos do conjunto à medida que forem sendo sorteados); métodos com sequência alternada de alocação; e outros³⁹.

Sigilo de alocação: o sigilo de alocação ocorre quando participantes e pesquisadores não conseguem saber ou prever em qual grupo os participantes serão alocados durante o recrutamento.

- Métodos considerados adequados: central de randomização (sistema por telefone, e-mail ou sistema interativo on-line); uso de embalagem idêntica de medicamentos sequencialmente numerados, preparados por uma farmácia independente; uso de envelopes sequencialmente numerados, lacrados e opacos.
- Métodos considerados inadequados: sequência de randomização postada abertamente em um quadro de avisos; uso de envelopes não lacrados ou transluzente; qualquer tipo de alocação previsível³⁹.

Desequilíbrios na linha de base: avalia o sucesso da randomização em produzir grupos de intervenção comparáveis em relação a importantes fatores prognósticos. Serão considerados um potencial viés, quando houver:

- Diferenças substanciais entre os tamanhos dos grupos em relação à razão de alocação pretendida.

- Desequilíbrios clinicamente relevantes em relação a um ou mais fatores prognósticos e valores dos desfechos na linha de base, que são improváveis de terem ocorrido ao acaso³⁹.

Desvios das intervenções pretendidas

O domínio contempla a avaliação do cegamento dos participantes e da equipe (cuidadores ou aqueles que estão administrando a intervenção); administração de intervenções adicionais que não estavam descritas no protocolo do estudo; falhas na implementação das intervenções pretendidas ou falta de adesão dos participantes às intervenções; utilização de análises apropriadas (análise por intenção de tratar, análise por intenção de tratar modificada ou análise por protocolo, entre outras)³⁹.

Cegamento: o cegamento dos participantes e da equipe diminui a probabilidade de que o conhecimento das intervenções possa afetar os resultados do estudo. Importante ressaltar que nem sempre é possível realizar o cegamento dos participantes e da equipe, como em estudos que comparam intervenções cirúrgicas com intervenções não cirúrgicas, mas o risco de viés não deve ser ignorado e continua presente em tais situações³⁹.



Análises apropriadas: quando o objetivo da RS for estimar o efeito da atribuição às intervenções (maioria das RS), é considerada análise apropriada aquela que segue os princípios da análise por intenção de tratar (ITT). Na análise por ITT, todos os pacientes randomizados devem ser incluídos na análise e analisados nos grupos em que foram randomizados, independentemente se realmente receberam a intervenção pretendida e se houve adesão por parte dos participantes²³.

A análise por ITT modificada, na qual participantes com dados faltantes são excluídos (vieses causados por dados faltantes são avaliados no domínio “dados incompletos dos desfechos”), também pode ser considerada adequada para esse domínio³⁹. Muitos estudos utilizam o termo ITT modificado para se referir à análise em que são incluídos apenas participantes que receberam pelo menos uma dose de tratamento. Esse tipo de análise não é considerado adequado. Também não são consideradas adequadas a análise por protocolo (que exclui participantes que não aderiram ao tratamento), a análise *as-treated* (participantes são analisados de acordo com a intervenção que eles realmente receberam, mesmo que eles tenham sido randomizados para outra intervenção), assim como a análise de casos completos (que exclui participantes com dados completos, ou seja, sem dados faltantes)³⁹.

Como discutido anteriormente, em alguns casos específicos, os autores podem ter como objetivo estimar o efeito da adesão às intervenções. Apesar de serem muito utilizadas nessas situações, as análises por protocolo simples e *as-treated* tendem a gerar estimativas bastante enviesadas e, conseqüentemente, maior risco de viés, quando comparado com análises por pro-

toloco com métodos estatísticos mais modernos que realizam ajustes para fatores prognósticos que predizem desvios das intervenções⁴⁰.

Dados faltantes dos desfechos

Dados faltantes dos desfechos em um ensaio clínico podem acontecer por inúmeras razões. Entre as mais comuns, podemos citar as seguintes: participantes decidem sair do estudo (desistências) ou não são mais encontrados (perda de seguimento); participantes não comparecem à visita em que os desfechos seriam mensurados; participantes não podem mais “experimentalmente” os desfechos (ex.: o participante morreu durante o estudo devido à intervenção ou não)³⁹.

A extensão do viés está diretamente relacionada à quantidade de dados faltantes. Porém não há limiar sensível quanto à porcentagem aceitável de dados faltantes. Existe um senso comum que de 5% a 10% sejam considerados uma taxa aceitável de dados faltantes, e mais de 20% uma taxa alta. Esse impacto do viés nas estimativas de efeito depende da taxa de dados faltantes, do tipo de desfecho, do risco dos eventos (no caso de desfechos dicotômicos) e na magnitude do efeito observado³⁹.

Não há viés na estimativa de efeito, apenas quando o desfecho é mensurado em todos ou quase todos (taxa muito pequena que provavelmente não possui influência) os participantes²³. A presença de dados faltantes dos desfechos é provável de causar viés quando há diferenças nas razões e na proporção da taxa de dados faltantes entre os grupos (indiretamente pode indicar que há relação entre o dado faltante e o verdadeiro valor do desfecho)³⁹.

Mensuração dos desfechos

A avaliação desse domínio envolve os seguintes aspectos: 1) utilização de métodos adequados para mensurar os desfechos (ex.: utilização de instrumentos validados ou não); 2) mensurações dos desfechos foram similares ou não entre os grupos; 3) quem é o verdadeiro avaliador do desfecho, o qual pode ser o próprio participante (desfechos relatados pelo participantes, como dor e função), o profissional da equipe (quando o desfecho é mensurado por uma avaliação clínica) ou um observador que não está envolvido diretamente na intervenção fornecida aos participantes; 4) cegamento do avaliador em relação à intervenção utilizada (cegamento assegurado e improvável de ter sido violado durante o estudo); e 5) a avaliação do desfecho é provável de ser influenciada pelo conhecimento da intervenção utilizada ou não (desfechos relatados pelos participantes ou que envolvem algum julgamento dos avaliadores, do profissional da equipe ou dos observadores são mais influenciáveis, ao contrário de desfechos que não envolvem julgamento, como mortalidade)³⁹.

Seleção dos resultados relatados

Esse domínio leva em conta os seguintes aspectos: 1) se o estudo foi analisado de acordo com o que foi planejado no protocolo, o qual foi finalizado antes da disponibilização dos dados para análise (revisores devem checar o protocolo, ou o registro do estudo, ou o plano de análise para checar se há inconsistências); 2) relato seletivo de apenas um ou de parte dos tempos de seguimento predefinidos, ou uso de múltiplos instrumentos para o mesmo desfecho, e relato seletivo dos dados coletados por um instrumento com resultado favorável; 3) relato seletivo de determinada análise (quando múltiplas análises foram realizadas), como o uso da análise mais favorável, quando tanto o resultado no tempo de seguimento e a mudança do resultado em relação à linha de base foram coletados, ou categorização de um desfecho que foi previamente contínuo baseado no resultado³⁹.

Os julgamentos de cada domínio vão servir como base para o julgamento geral do risco de viés de um ECR para cada desfecho específico, de acordo com os critérios apresentados no **Quadro 8**.

Quadro 8 – Critérios para julgamento geral do risco de viés de ECR para um desfecho específico

Julgamento geral do risco de viés	Critérios
Baixo risco de viés	Se todos os domínios forem julgados como “baixo risco de viés”.
Algumas preocupações	Se pelo menos um domínio for julgado como “algumas preocupações”, e nenhum domínio for julgado como “alto risco de viés”.
Alto risco de viés	Se pelo menos um domínio for julgado como “alto risco”; ou se múltiplos domínios forem julgados como “algumas preocupações”, de forma que a confiança no resultado é reduzida substancialmente.

Fonte: Ferramenta Cochrane RoB-2³⁹.

3.5.7.2 Como realizar a avaliação do risco de viés?

Dois avaliadores independentes devem realizar a avaliação do risco de viés dos estudos incluídos, a fim de minimizar o risco de julgamentos equivocados e influenciados pelo viés dos próprios autores da RS. Os autores da RS também devem informar se as discordâncias serão resolvidas por meio de consenso ou se um terceiro avaliador realizará a decisão, solucionando-as.

A avaliação do risco de viés deve ser realizada em nível dos principais desfechos da RS (geralmente os desfechos selecionados para serem apresentados na tabela sumária dos resultados), em cada tempo de seguimento avaliado. O julgamento dos domínios pode ser influenciado pelo tipo de desfecho (por exemplo, a abordagem para avaliar o viés devido à mensuração de um desfecho relatado pelo paciente não é a mesma para avaliar a mensuração de mortalidade). Alguns domínios (por exemplo, viés proveniente do processo de randomização) podem afetar o estudo como um todo³⁶.



Para a realização da avaliação do risco de viés, os revisores devem coletar informações de diversas fontes dos estudos, e não apenas da publicação principal. Informações dos protocolos dos estudos, registros de ensaios clínicos, documentos das companhias farmacêuticas, agências regulatórias devem ser utilizados para uma avaliação mais completa do risco de viés³⁶.

3.5.7.3 Incorporação da avaliação do risco de viés nas análises

É de extrema importância que os autores da RS incorporem a avaliação do risco de viés na análise e na interpretação dos achados, e não ignorem os vieses identificados na avaliação. Para isso, os autores devem identificar e justificar se vão utilizar apenas os estudos julgados como “baixo risco de viés” ou se todos os estudos (independentemente do julgamento do risco de viés) serão utilizados nas conclusões. Essa decisão é dos autores da RS, que devem considerar o balanço entre viés e precisão, ou seja, o quanto se ganha (resultados sem vieses) e se perde com a exclusão de estudos julgados como “alto risco de viés” (perda de precisão, por incluir menos estudos na meta-análise, o que normalmente gera intervalos de confiança mais amplos)³⁶.

Caso os autores de RS optem por utilizar apenas estudos com “baixo risco de viés” na análise primária, é recomendado que análises de sensibilidade sejam realizadas para investigar como os resultados podem ser influenciados se os estudos julgados como “alto risco de viés” também forem incluídos nas análises³⁶.

A avaliação do risco de viés também é um dos domínios incorporados na avaliação da qualidade da evidência pela abordagem Grading of Recommendations Assessment, Development and Evaluation (GRADE), o que reforça a importância e o impacto dos possíveis vieses metodológicos na síntese dos resultados da RS.

3.5.8 | Síntese dos resultados

Uma etapa importante de uma RS é a decisão de realizar ou não uma síntese quantitativa dos resultados de dois ou mais estudos por meio de uma meta-análise⁴. Existem algumas situações em que não é recomendado ou não é possível realizar uma meta-análise: 1) presença de diversidade clínica e metodológica substancial entre os estudos identificados; 2) existência apenas de um ECR incluído para determinada comparação e determinado desfecho; 3) dados disponíveis dos desfechos são insuficientes (por exemplo, valores de média e desvio-padrão não foram relatados nos grupos de intervenção), impossibilitando o cálculo (e até mesmo a imputação de alguns dados) das estimativas de efeito. Caso não seja possível ou apropriado realizar uma meta-análise, uma síntese narrativa deverá ser feita.

3.5.8.1 Síntese qualitativa

Apesar de ser comumente utilizada, o uso de síntese qualitativa é bastante criticado pela falta de transparência e pelo potencial viés que pode gerar nas conclusões das RS⁴¹. Os resultados normalmente são interpretados de forma subjetiva, sem que haja justificativa para a ênfase que é dada para determinados resultados⁴².

Os autores da RS devem estruturar, de forma consistente, o relato dos resultados em uma síntese qualitativa, priorizando a apresentação dos dados de forma quantitativa e sempre ponderando o risco de viés dos estudos. Sempre que possível, as estimativas de efeito devem ser apresentadas com suas medidas de dispersão, seja em formato de gráficos ou nas tabelas e no texto, e agrupadas de acordo com as comparações e desfechos avaliados.

Em situações em que informações das estimativas de efeito não estão disponíveis, os autores devem relatar o valor-p (se disponível) e/ou a direção de efeito de cada estudo. Quando os autores da RS tiverem disponíveis apenas o valor-p e a direção do efeito, métodos de combinação dos valores-p podem ser utilizados para sintetizar os resultados⁴³. Esse método é útil para responder se há evidência de que existe diferença de efeito entre os grupos de comparação em pelo menos um estudo, e pode ser apresentado pelo gráfico de Albatross⁴⁴. Quando apenas a direção do efeito estiver disponível, os autores da RS podem utilizar o método de contagem de votos baseados na direção do efeito⁴⁵. Esse método compara o número de estudos que demonstrou benefício ou dano para cada desfecho e pode ser apresentado pelo gráfico de Harvest⁴⁶. É importante destacar que esses métodos possuem sérias limitações e são pouco utilizados nas RS, porém podem ser mais úteis do que uma simples descrição qualitativa dos estudos incluídos⁴⁵.



As estimativas de efeito devem ser apresentadas com suas medidas de dispersão.

3.5.8.2 Meta-análise

Meta-análises geram estimativas de efeito sumárias (com o IC) que sintetizam os efeitos de uma intervenção experimental comparada como uma intervenção comparadora⁴⁷. O uso da meta-análise não garante que os resultados da RS sejam válidos e, assim como outros métodos estatísticos, seu uso pode ser inadequado.

Meta-análises só devem ser realizadas quando os estudos são considerados suficientemente homogêneos do ponto de vista clínico e metodológico^{4, 47}. O julgamento sobre os estudos serem homogêneos o bastante e, portanto, a combinação dos resultados desses estudos em uma meta-análise ser adequada devem ser baseados no entendimento da pergunta de pesquisa, características dos estudos e da interpretação dos resultados^{4, 47}. A heterogeneidade estatística não deve ser levada em conta na decisão de realizar ou não meta-análise⁴⁷. Caso não seja possível ou apropriado realizá-la, uma síntese qualitativa deverá ser executada.

Quando apropriada, a realização de meta-análises em uma RS possui as seguintes vantagens⁴⁷: 1) aumento da precisão das estimativas de efeitos (quando comparada com estudos individuais), já que geralmente a precisão está diretamente relacionada com a quantidade de informação disponível; 2) solução de questões que não conseguem ser respondidas por estudos individuais, já que normalmente os estudos incluem apenas um tipo específico de participantes e intervenções; e 3) solução de controvérsias que surgem quando os estudos individuais apresentam resultados conflitantes (as razões para resultados diferentes devem ser exploradas na RS).

Princípios da meta-análise

A meta-análise é um processo que ocorre em dois estágios. No primeiro momento, uma estimativa de efeito (com o erro-padrão) é calculada para cada estudo, a fim de descrever os efeitos das intervenções^{8, 47}. As estimativas de efeito (como já descritas no tópico 4.5.6) podem ser RR ou OR para desfechos dicotômicos; e DM ou DMP para desfechos contínuos; ou, ainda, *hazard ratio* para desfecho do tipo tempo para o evento. Em um segundo momento, uma estimativa sumária de efeito é calculada por meio de uma média ponderada (cada estudo terá um peso) das estimativas de efeito individuais dos estudos. A média ponderada pode ser defendida como^{8, 47}:

$$\text{Média Ponderada} = \frac{\text{soma das (estimativa de cada estudo x peso de cada estudo)}}{\text{soma dos pesos}} = \frac{\sum Y_i W_i}{\sum W_i}$$

Na fórmula da média ponderada, Y_i representa a estimativa de efeito de cada estudo, W_i é o peso dado para cada estudo na meta-análise, e \sum é a soma entre todos os estudos incluídos na meta-análise. Quanto maior o peso dado a um estudo na meta-análise, maior será a contribuição dele na média ponderada e, conseqüentemente, na estimativa sumária final. Normalmente, é dado maior peso para os estudos com maior precisão (geralmente estudos com maiores amostras), mas alguns ajustes são realizados a depender do modelo de meta-análise adotado^{8, 47}.

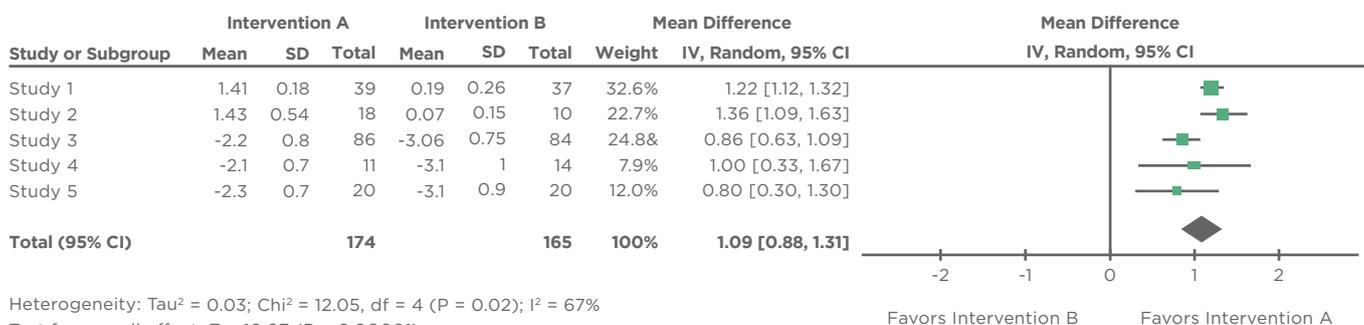
Além da estimativa sumária de efeito, é necessário calcular o erro-padrão dessa estimativa sumária, o qual é usado para gerar o IC (normalmente com o nível de confiança de 95%) da estimativa sumária (diz respeito à precisão da estimativa) e o valor p (diz respeito a se a hipótese nula de nenhum efeito da intervenção foi rejeitada ou não)^{8, 47}.

Todos os métodos de meta-análise devem incorporar uma medida para avaliação da heterogeneidade estatística (variabilidade) das estimativas de efeitos entre os estudos incluídos. Essa avaliação é importante para entender se a variação das estimativas entre os estudos ocorreu pelo acaso ou se é grande o suficiente para indicar inconsistência entre eles^{8, 47}.

Apresentação da meta-análise

As meta-análises devem ser apresentadas usando o gráfico da floresta (ou *forest plot*, em inglês) (**Figura 1**). No gráfico, são apresentadas as estimativas (com o IC) de cada estudo individual, assim como da estimativa sumária e de seu respectivo IC. Cada estudo é representado no gráfico por um quadrado (estimativa de efeito) com uma linha horizontal (IC) que se estende de um lado a outro dele. Quanto maior a área do quadrado, maior o peso (geralmente, os estudos com intervalos de confiança são mais estreitos) do estudo na meta-análise^{8, 47}. A estimativa de efeito sumária com o IC é representada pelo diamante (losango) na parte inferior do gráfico. O local do diamante representa a estimativa pontual do tamanho do efeito, e sua extensão representa o IC (precisão) associado a essa estimativa^{8, 47}.

Figura 1 - Exemplo de *forest plot* com desfecho contínuo



Fonte: Elaboração própria.

A linha vertical que cruza todo o gráfico representa o efeito-nulo (significância estatística) entre as intervenções comparadas. Essa linha está localizada no 1, para desfechos dicotômicos, ou no 0, para desfechos contínuos. Quando os estudos individuais, representados pelos quadrados e suas respectivas linhas horizontais cruzam a linha vertical, podemos afirmar que a hipótese nula não foi rejeitada, ou seja, as diferenças entre os efeitos das intervenções não são estatisticamente significantes^{8, 47}. O mesmo raciocínio é aplicado para a interpretação da posição do diamante.

Modelos estatísticos de meta-análise

A maioria das meta-análises é baseada no uso de um dos dois modelos estatísticos: modelo de efeito fixo e modelo de efeitos aleatórios. Os autores da RS devem indicar qual modelo foi utilizado e a justificativa da escolha dele no texto. Nesse tópico, iremos apresentar os conceitos de cada modelo, abordar as diferenças entre eles, suas indicações e implicações na condução da RS.

Modelo de efeito fixo

No modelo de efeito fixo, supõe-se que existe um efeito de intervenção verdadeiro (um efeito comum) subjacente a todos os estudos na análise, e que as diferenças nas estimativas observadas entre os estudos são explicadas exclusivamente pelo erro amostral (aleatório)^{8, 47}. O efeito de intervenção verdadeiro representa a estimativa de efeito observada em um estudo hipotético com tamanho amostral infinito, portanto sem erro amostral⁸. Sob esse modelo, é assumido que todos os estudos na meta-análise compartilham um efeito verdadeiro comum. Todos os fatores que poderiam influenciar a estimativa de efeito são iguais em todos os estudos⁸.

Nesse modelo, como a variação das estimativas de efeito entre os estudos são explicadas apenas pelo erro aleatório, pode-se assumir que, se cada estudo tivesse um tamanho amostral infinito (portanto o erro amostral seria zero), o efeito observado de cada estudo seria igual ao efeito verdadeiro⁸. O peso dado a cada estudo na meta-análise normalmente é computado pelo inverso da variância do estudo. Isso é realizado com o objetivo de obter uma estimativa de efeito mais precisa, minimizando o impacto da variância (devido ao erro amostral) de cada estudo⁸.

Modelo de efeitos aleatórios

Na maioria das RS, o conceito do modelo de efeito fixo, em que todos os estudos compartilham um efeito de tratamento verdadeiro, não é plausível⁸. Quando ocorre o julgamento de que a realização de uma meta-análise é recomendada, é porque os estudos são suficientemente homogêneos na perspectiva clínica e metodológica. Entretanto, dificilmente assume-se que esses estudos são idênticos para justificar a suposição de que o efeito de tratamento verdadeiro é o mesmo em todos eles^{4, 8}.

Estudos provavelmente irão diferir em relação a algumas características dos participantes, à forma como os tratamentos são realizados, entre outros pontos. É esperado então que haja diferentes estimativas de efeito entre estudos diferentes. Uma maneira de incorporar essa variação em uma meta-análise é pelo uso do modelo de efeitos aleatórios^{8, 47}. No modelo de efeitos aleatórios, é incorporada a suposição de que o efeito de intervenção verdadeiro pode variar de estudo para estudo, e essa variação não pode ser explicada apenas pelo erro amostral^{8, 47}. Nesse modelo, não há um efeito verdadeiro comum (como no modelo de efeito fixo) e, se fosse possível realizar um número infinito de estudos, os efeitos verdadeiros desses estudos teriam distribuição normal sobre alguma média^{8, 47}.

Para exemplificar essa suposição de que estudos diferentes produzem estimativas de efeito de tratamento diferentes, pode-se imaginar que a estimativa pode ser menor ou maior em estudos nos quais os participantes têm idade mais avançada, ou têm um nível socioeducacional maior, ou ainda uma condição de saúde melhor que participantes de outros estudos⁸. Variações nas características da intervenção também podem causar estimativas de efeitos maiores ou menores.

Em suma, no modelo de efeitos aleatórios, a estimativa sumária na meta-análise nada mais é do que uma média das estimativas de efeitos dos estudos individuais, sendo que os estudos com maior tamanho amostral (os quais possuem menor erro-padrão) terão maior peso na meta-análise^{8, 47}. Assim como no modelo de efeito fixo, o peso dado a cada estudo é calculado pelo inverso da variância. A diferença é que, além de levar em conta a variabilidade das estimativas de cada estudo causada pelo erro amostral, é necessário também computar a variância entre as estimativas dos estudos, conhecida como *tau-squared* (T^2)⁸.

Efeito fixo versus efeitos aleatórios

No modelo de efeito fixo, os estudos pequenos têm pouca influência sobre a meta-análise, já que há melhor informação sobre a mesma estimativa de efeito em estudos maiores. Já no modelo de efeitos aleatórios, como o objetivo é ter uma estimativa de efeito médio dos estudos, cada estudo é responsável por uma estimativa de efeito diferente⁸. Assim, todas essas estimativas diferentes, no modelo de efeitos aleatórios, estão contempladas na estimativa sumária. Na presença de variabilidade entre os estudos (T^2 diferente de zero), o peso dado a eles é mais balanceado no modelo de efeitos randômicos do que no modelo de efeito fixo. Ou seja, no modelo de efeitos aleatórios, é dado relativamente maior peso aos estudos menores e relativamente menor peso aos estudos maiores^{8, 47}.

Os intervalos de confiança ao redor das estimativas de efeito são mais largos no modelo de efeitos aleatórios do que no modelo de efeito fixo^{8, 47}. Já na ausência de heterogeneidade, resultados idênticos serão encontrados nos dois modelos.

A decisão de qual modelo de meta-análise utilizar não deve ser realizada levando-se em conta o resultado da avaliação da heterogeneidade estatís-

tica; por exemplo, utilizando-se níveis de corte da medida I^2 (discutido nos próximos tópicos). A escolha do método deve ser realizada tendo em vista apenas a heterogeneidade clínica e metodológica dos estudos⁴⁷. Como já explicado, o modelo de efeito fixo deve ser escolhido quando se acredita que todos os estudos incluídos são funcionalmente idênticos, e o objetivo da meta-análise é gerar uma estimativa de efeito comum para a população identificada, não generalizar essa estimativa para outras populações⁴⁷. Essa situação costuma acontecer, por exemplo, quando a RS inclui estudos que avaliam o uso de algum medicamento (ministrado da mesma forma) em pacientes com alguma condição (e que apresentam características similares) do mesmo grupo de pesquisa ou fontes de financiamento. Entretanto, isso é bastante incomum na maioria das RS.

A escolha pelo modelo de efeitos aleatórios deve ser realizada nas situações em que os autores da RS estão sumarizando dados de vários estudos que foram realizados independentemente, e provavelmente não são funcionalmente idênticos⁴⁷. O objetivo dessa análise geralmente é generalizar os achados para cenários variados, e não somente para os estudos incluídos na meta-análise.

Métodos estatísticos de meta-análise

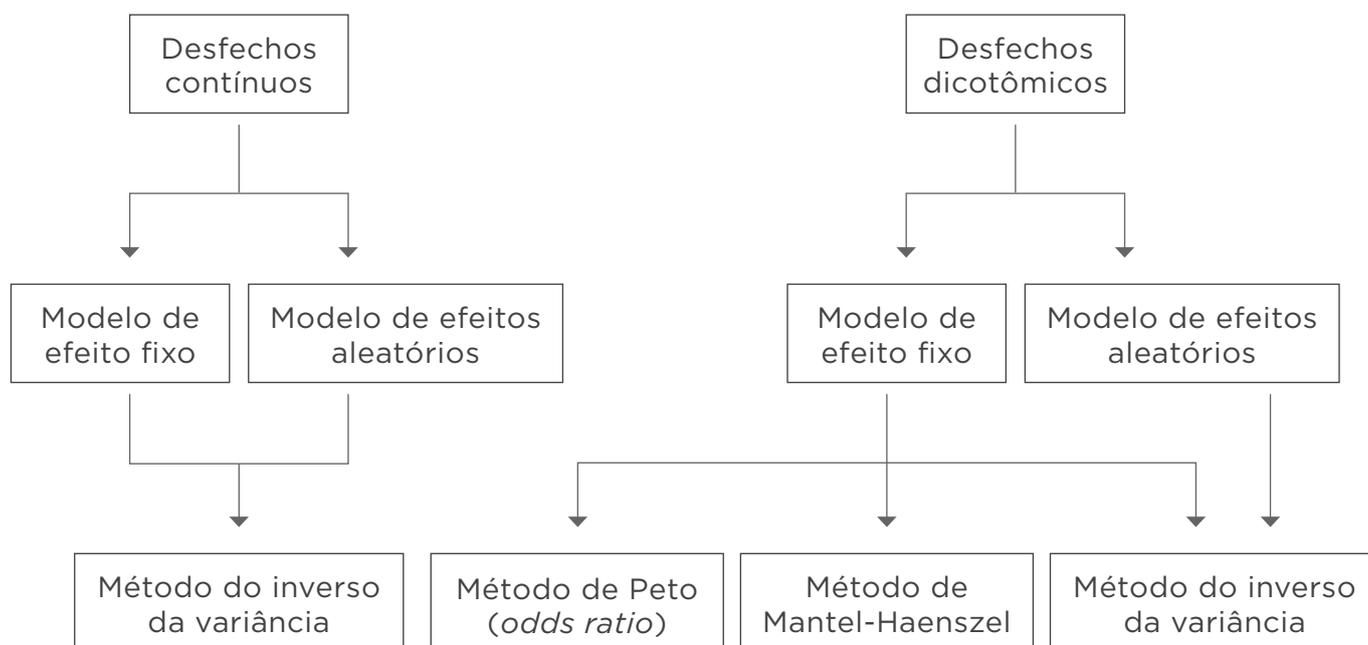
Além de decidir o modelo a ser utilizado na meta-análise, os autores de RS precisam definir qual método estatístico será implementado nela. O modelo mais simples e mais utilizado nas meta-análises é o método do inverso da variância, o qual pode ser usado para todas as métricas dos desfechos dicotômicos (RR, OR etc.) e desfechos contínuos (DM e DMP)⁴. Como o próprio nome diz, nesse método o peso dado para cada estudo é calculado pelo inverso da variância da estimativa de efeito. Dessa forma, estudos maiores, os quais possuem menores erros-padrão, recebem maior peso do que os estudos menores. Esse processo minimiza a imprecisão na estimativa de efeito agrupada⁸.

O método do inverso da variância é utilizado para o modelo de efeito fixo, mas há uma variação desse método, conhecido como método de Dersimonian e Laird, que é utilizada para o modelo de efeitos aleatórios⁸. Esse método é usado na forma mais básica pelo programa gratuito Review Manager, assim como pela maioria dos programas estatísticos que rodam meta-análise (Stata, R, Comprehensive Meta-Analysis etc.). Para utilizar esse método, os autores da RS normalmente devem fornecer dados no formato da tabela 2x2 (eventos, não eventos e tamanho de amostra de cada estudo) e média, desvio-padrão e tamanho amostral de cada estudo para desfechos contínuos⁸.

Outros dois métodos estatísticos podem ser usados para meta-análise de desfechos dicotômicos pelo modelo de efeito fixo: Mantel-Haenszel e Peto. O método de Mantel-Haenszel pode ser utilizado com qualquer métrica de desfecho dicotômico (RR, OR etc.) e é indicado por ter melhores propriedades estatísticas (comparado com os outros métodos) quando os estudos possuem amostra pequena e há poucos eventos na meta-análise⁸.

Já o método de Peto pode ser usado apenas combinando estimativas em OR e com modelo de efeito fixo⁸. Esse método usa a abordagem do inverso da variância, mas utiliza um método aproximado de estimar as medidas em log OR, assim como diferentes pesos para cada estudo. Esse método é indicado quando as estimativas de efeito são pequenas (OR próximos a 1), eventos não são comuns ou quando não há eventos em um dos grupos e os estudos têm amostras similares nos grupos de tratamento⁸. Como isso é incomum, esse método é pouco utilizado. A **Figura 2** resume os principais métodos empregados para meta-análise.

Figura 2 – Principais métodos estatísticos para meta-análise de desfechos contínuos e dicotômicos



Fonte: Adaptado da primeira edição das *Diretrizes Metodológicas: elaboração de revisão sistemática e meta-análise de ensaios clínicos randomizados*⁴⁸.

Meta-análise de desfechos dicotômicos

Meta-análises de desfechos dicotômicos podem ser expressas em medidas relativas (ex.: OR e RR) ou em medidas absolutas (ex.: diferença de risco e NNT). Geralmente, é recomendado que as meta-análises sejam apresentadas usando RR ou OR, já que medidas absolutas podem gerar estimativas de efeito pouco consistentes em algumas situações⁸. Apesar de tanto o RR como OR serem medidas válidas para descrever um efeito da intervenção, elas podem ser bem diferentes quando a taxa de eventos é alta. Nessas situações, o OR tende a superestimar o efeito da intervenção. Os valores de OR e RR tendem a se aproximar quando os eventos são incomuns. Uma crítica ao uso de OR é a dificuldade de interpretação e a aplicabilidade dos resultados, principalmente por parte dos clínicos⁸.

Quando os autores da RS optarem por utilizar medidas de efeito relativa, é recomendado que os resultados também sejam apresentados em uma maneira mais fácil de interpretar, como as medidas de efeito absoluto (diferenças de risco e NNT)⁸. Isso pode ser realizado na tabela com o sumário dos achados, que será apresentada no tópico 4.6.5 destas diretrizes.

Meta-análise de desfechos contínuos

Para a realização de meta-análise de desfechos contínuos, os autores da RS deverão optar por expressar os resultados geralmente em DM ou DMP. Quando os estudos relatam o desfecho na mesma escala, a DM é normalmente usada. Já quando os estudos usam diferentes escalas, é recomendado utilizar a DMP^{8, 47}.

Em alguns ECR, é comum os resultados estarem apresentados como mudança em relação à linha de base em vez de valores pós-tratamento. Combinar valores de pós-tratamento e valores de mudança em relação à linha de base em uma meta-análise é possível apenas quando os resultados foram apresentados em diferença de média. Do ponto de vista estatístico, não é possível fazer essa combinação quando os resultados forem apresentados em diferença de média padronizada⁴⁷.

3.5.9 | Avaliação da heterogeneidade

Heterogeneidade pode ser definida como qualquer tipo de variabilidade entre os estudos incluídos em uma RS. Os estudos sempre irão variar, de alguma forma, entre eles, seja em relação a características da população, a intervenções e a desfechos (diversidade clínica), seja em relação aos aspectos metodológicos, como presença de vieses (diversidade metodológica)^{47, 49}. Uma consequência da presença de diversidade clínica e/ou metodológica é a variabilidade nas estimativas de efeito de tratamento entre os estudos, que é chamada de heterogeneidade estatística (ou simplesmente heterogeneidade). A heterogeneidade estatística ocorre quando as estimativas de efeitos dos estudos diferem entre eles mais do que se esperaria apenas pelo acaso (erro aleatório)^{8, 47}. A investigação da presença da heterogeneidade (estatística) é um processo importante na realização das RS, sendo essencial considerar o quanto os resultados dos estudos são consistentes entre eles.

3.5.9.1 Identificação e quantificação da heterogeneidade estatística

Uma forma visual e simples de os autores identificarem a presença de heterogeneidade na meta-análise é observar se não há sobreposição (ou pouca sobreposição) dos intervalos de confiança dos resultados dos estudos incluídos⁴⁷. A heterogeneidade também deve ser avaliada por medidas estatísticas, como o teste do chi-quadrado (Chi^2), valores de I^2 e variância entre os estudos, que é denominada de tau-quadrado (T^2) (todas essas medidas estão disponíveis nos *forest plots* do software Review Manager).

O teste do Chi^2 é baseado na estatística Q, no qual a hipótese nula testada é que os estudos incluídos na meta-análise não apresentam variabilidade nas estimativas de efeitos (além da esperada pelo erro aleatório proveniente dos tamanhos amostrais)⁸. A estatística Q segue uma distribuição Chi^2 com graus de liberdade iguais a $k - 1$ (em que k é o número de estudos incluídos na meta-análise), gerando um p-valor correspondente⁸. Um p-valor baixo (ou um valor alto de Q em relação aos graus de liberdade) indica evidência de heterogeneidade estatística. O teste do Chi^2 possui baixo poder estatístico (principalmente quando há poucos estudos na meta-análise, ou quando os estudos possuem tamanho amostral pequeno) e devem ser interpretados com precaução⁴⁷. Enquanto um resultado estatisticamente significativo ($p < 0,1$) pode evidenciar a presença de heterogeneidade estatística; um resultado não significativo ($p \geq 0,1$) não deve ser interpretado como ausência de heterogeneidade^{8, 47}.

Tão importante quanto investigar se há heterogeneidade ou não na meta-análise (já que alguma heterogeneidade sempre vai existir), é a avaliação do seu impacto nos resultados por parte dos autores de RS. Isso é realizado pela medida estatística I^2 , que descreve em porcentagem (varia de 0% a 100%) o quanto da variabilidade observada nas estimativas de efeito de tratamento pode ser explicada pela heterogeneidade⁴⁹. Valores de I^2 próximos de 0% indicam que quase toda a variabilidade observada é devido ao erro aleatório (explicado por diferenças no tamanho amostral), portanto não há heterogeneidade observada⁴⁹. Existem alguns limites propostos para interpretar os valores de I^2 , mas essa interpretação deve ser realizada com cuidado (há incerteza dos valores I^2 quando o número de estudos é pequeno). Uma interpretação simples sugerida é⁴⁷:

- **0% a 40%:** pode não ser importante.
- **30% a 60%:** pode representar heterogeneidade moderada.
- **50% a 90%:** pode representar heterogeneidade substancial.
- **75% a 100%:** heterogeneidade considerável.

Outra medida importante na avaliação da heterogeneidade no contexto da meta-análise pelo modelo de efeitos aleatórios é o T^2 , que reflete a extensão da variação dos efeitos de tratamento observados nos diferentes estudos (reflete a variação dos tamanhos de efeito verdadeiros)⁸. Pelo fato de o T^2 refletir uma variância, ele não é uma medida muito intuitiva. O valor de T^2 pode ser ≥ 0 , sendo que valores mais altos indicam uma variância maior das estimativas de efeitos de tratamento; portanto uma heterogeneidade estatística maior⁸.



É importante que a abordagem da heterogeneidade no contexto da meta-análise incorpore tanto uma medida da sua magnitude, como uma medida da incerteza. Esta atualização da diretriz recomenda que a medida da magnitude pode ser representada pelo valor de T^2 ou I^2 (ou ambos), enquanto a incerteza sobre a presença de uma aparente heterogeneidade pode ser expressa usando o p-valor do teste Chi^2 ⁸.

3.5.9.2 Como lidar com a heterogeneidade em uma revisão sistemática

A avaliação da heterogeneidade por si só não é importante, caso os autores não levem em conta seu impacto nos resultados encontrados. A heterogeneidade afeta o quanto as conclusões podem ser generalizadas⁴. Dessa forma, os autores de RS sempre devem adotar estratégias para lidar com a presença da heterogeneidade (independentemente se o modelo de meta-análise adotado foi de efeito fixo ou efeitos aleatórios). Importante destacar que, no modelo de efeito fixo, os intervalos de confiança ignoram a extensão da heterogeneidade, enquanto o modelo de efeitos aleatórios pode ser usado para incorporar a heterogeneidade entre os estudos⁴⁷.

O primeiro passo, quando na presença de heterogeneidade, é checar novamente se não houve erros no momento de extração dos dados ou de inclusão deles no software da meta-análise (ex.: dados de erro-padrão podem ter sido incluídos erroneamente como desvio-padrão)⁴⁷. O próximo passo é explorar a heterogeneidade, na tentativa de descobrir potenciais causas. Análise de subgrupo e/ou de metarregressão é alternativa útil para explorar a heterogeneidade. Porém essas análises precisam ser definidas a priori no protocolo ou no registro da RS. Detalhes sobre análise de subgrupo e de metarregressão serão detalhados no tópico a seguir.

Outra estratégia para lidar com a heterogeneidade é excluir da meta-análise estudos considerados *outliers* (geralmente um ou mais estudos podem apresentar resultados totalmente conflitantes do resto dos estudos) da meta-análise. Essa decisão deve ser feita com muita cautela, pois podem introduzir viés na RS, e só devem ser tomadas se há uma explicação óbvia para esses estudos serem considerados *outliers*⁴⁷.

3.5.10 | Análises de subgrupo

Análises de subgrupo envolvem dividir todos os dados dos participantes na meta-análise em subgrupos, geralmente de acordo com determinadas características dos participantes (sexo, idade etc.), características dos estudos incluídos (diferentes localizações geográficas, por exemplo), características da intervenção (dose, por exemplo) ou aspectos metodológicos dos estudos (processo de randomização)^{47, 50}. As análises de subgrupo podem ser realizadas como uma forma de investigar a heterogeneidade dos resultados (como explicado no tópico anterior), ou para responder questões específicas de subgrupos de pacientes ou tipos de intervenção^{47, 50}.

É importante destacar que essas análises são de natureza exploratória (muito úteis e importantes para gerar hipóteses), não sendo baseadas em comparações randomizadas; portanto conclusões definitivas não podem ser tomadas a partir delas⁴⁷. É muito comum análises de subgrupos apresentarem resultados falso-positivos, e a chance de um falso-positivo aumenta na mesma medida que o número de análises realizadas⁴⁷.

3.5.10.1 Princípios para realização de análise de subgrupo

Para a realização de análises de subgrupo, os autores de RS devem ter cautela na realização e na interpretação dos resultados⁴. As características a serem investigadas nas análises de subgrupo devem ser definidas a priori no registro e/ou no protocolo da RS. A pré-especificação das análises reduz a probabilidade de achados espúrios, limitando o número de análises a serem realizadas e evitando que o conhecimento dos resultados da RS influencie a decisão de quais características devam ser analisadas⁴⁷. Vale ressaltar que podem ocorrer situações em que determinada característica não foi especificada a priori, mas se apresen-

ta claramente importante para ser investigada. Os autores da RS devem explorar determinada característica, mas deixando claro se tratar de uma análise de subgrupo *post-hoc*.

Para a realização de análises de subgrupo, é necessário haver número suficiente de estudos incluídos. Recomenda-se que, para cada característica investigada, deve haver pelo menos dez estudos incluídos na meta-análise. Além disso, é importante que o número de estudos seja balanceado em cada subgrupo. Portanto um número baixo de características deve ser planejado para análise de subgrupos, já que, quanto maior o número de análises, maior a chance de resultados falso-positivos^{47, 50}. Não existe uma regra em relação ao número de características, já que varia de acordo com o número de estudos incluídos em cada meta-análise, mas normalmente não mais que três características devem ser exploradas. A definição (a priori) das características a serem investigadas nas análises de subgrupo deve ser motivada por hipóteses clínicas e, de preferência, baseada em evidências de estudos prévios⁴⁷. Essas características devem ter capacidade de modificar o efeito das intervenções, ou influenciar como uma intervenção pode afetar o desfecho⁴⁷. Cada análise de subgrupo planejada deve vir acompanhada de uma justificativa para a sua execução, incluindo a direção do efeito esperado nos subgrupos planejados (aumento ou redução do efeito da intervenção).

3.5.10.2 Condução e interpretação de análises de subgrupo

Autores de RS não devem interpretar que a heterogeneidade entre os estudos é explicada pela análise de subgrupo, simplesmente comparando a significância estatística dos resultados de cada subgrupo (um subgrupo apresenta um efeito estatisticamente significativo e o outro subgrupo não). O mesmo vale para a magnitude do efeito encontrado em cada subgrupo⁴⁷.

As análises de subgrupo podem ser feitas na maioria dos softwares que realizam meta-análise (como o RevMan, Stata, R, entre outros), por meio de um teste de significância estatística simples que avalia diferenças entre dois ou mais subgrupos⁸. Análises de subgrupo podem ser realizadas tanto usando o modelo de efeito fixo como o modelo de efeitos aleatórios. Um efeito de subgrupo estatisticamente significativo (normalmente é recomendado utilizar um p-valor menor que 0,05) significa que a característica explorada modifica o efeito do tratamento⁵⁰. É importante observar se o efeito de subgrupo é qualitativo (subgrupos apresentam direção de efeitos opostos) ou quantitativo (o tamanho do efeito varia, mas a direção não)^{47, 50}.

Além do teste que avalia a diferença entre os grupos, autores de RS devem observar a extensão da heterogeneidade estatística em cada subgrupo. Caso a heterogeneidade ainda seja importante em cada subgrupo, pode-se supor que a característica investigada na análise de subgrupo não explica a heterogeneidade como um todo⁵⁰.

3.5.11 | Metarregressão

Análises de metarregressão podem ser realizadas como alternativa para investigar diferenças entre subgrupos, quando uma meta-análise com modelo de efeitos aleatórios é utilizada e há disponibilidade de pelo menos dez estudos no desfecho em análise. Além disso, esse tipo de análise pode investigar características (variáveis) contínuas, assim como variáveis categóricas, permitindo inclusive que as características sejam investigadas simultaneamente^{8, 47}. A análise de metarregressão é muito similar a uma análise de regressão simples, sendo o efeito da intervenção a variável desfecho e as características investigadas as variáveis explanatórias ou covariáveis. O resultado da análise de metarregressão gera um coeficiente de regressão, o qual explica como

a(s) covariável(eis) influencia(m) o efeito da intervenção^{8, 47}. Um coeficiente de regressão estatisticamente significativo indica que há uma relação linear entre o efeito da intervenção e a covariável estudada⁴⁷.

Os princípios e pré-requisitos para a realização de análise de metarregressão em uma RS são os mesmos da análise de subgrupo, como a pré-especificação e a escolha das covariáveis motivadas por hipóteses clínicas e embasadas por evidência prévias. É importante destacar que não é possível realizar análise de metarregressão no programa Review Manager, mas é possível realizar em outros programas, como R e Stata).

3.5.12 | Análises de sensibilidade

É recomendado que autores de RS realizem análises de sensibilidade para avaliar o quão robustos são os seus achados⁴. Idealmente, as análises de sensibilidade devem ser especificadas, a priori, no registro ou protocolo da RS. Entretanto, muitas situações só são identificadas no decorrer da revisão, não sendo possível tal planejamento.

Análises de sensibilidade são análises alternativas da análise principal ou de meta-análises, em que decisões sobre determinados aspectos são realizadas de forma diferente para avaliar se os achados da revisão são robustos frente às decisões (como o método de imputação de dados adotado, inclusão de estudos com alto risco de viés e outras decisões limítrofes) tomadas pelos autores⁴⁷. Quando os resultados da análise de sensibilidade são similares aos da análise principal, ou seja, decisões diferentes de determinados aspectos não afetaram o resultado geral, há o aumento da confiança nos resultados encontrados na RS. Caso as análises de sensibilidade apresentem resultados distintos da análise principal, os achados da RS devem ser interpretados com cautela⁴.

Inicialmente, é importante destacar que análises de sensibilidade não são similares a análises de subgrupo, apesar de autores de RS se confundirem quanto à utilização desses termos. Isso ocorre porque, na análise de sensibilidade, assim como na análise de subgrupo, ela também é restrita a uma parte da totalidade dos estudos incluídos. Elas se diferenciam, porém, porque análises de sensibilidade não realizam comparações das estimativas de efeitos da intervenção dos subgrupos, e sim comparações informais entre diferentes formas de estimar a mesma coisa⁴⁷.

Apesar de não haver regras sobre quais decisões devem ser testadas nas análises de sensibilidade, as situações listadas a seguir são exemplos comumente analisados⁴⁷: 1) resumos de congressos ou conferências (em que não há posterior publicação completa) devem ser incluídos na revisão?; 2) características dos participantes: estudos em que a maioria, mas não toda a população incluída, preenche todos os critérios de inclusão da RS devem ser incluídos?; 3) características da intervenção: quais faixas de doses da intervenção devem ser incluídas na meta-análise?; 4) características do grupo comparador, em que definições diferentes são usadas para se referir a “cuidado usual”; 5) desenho do estudo: estudos com cegamento inadequado de participantes e/ou avaliadores devem ser incluídos?; 6) imputação de dados: como devem ser imputados os dados faltantes?; 7) análise dos dados: devem ser utilizados dados de pós-intervenção ou mudança em relação a linha de base?; e 8) modelo de meta-análise usado: deve ser utilizado modelo de efeito fixo ou efeitos aleatórios?.

3.5.13 | Avaliação do viés de publicação

Autores de RS devem sempre considerar o impacto do viés de publicação nos resultados, já que recomendações para prática clínica de intervenções ineficazes e prejudiciais podem ocorrer devido a esse viés, que tende a superestimar as estimativas de efeitos das intervenções. Viés de publicação pode ser definido como a publicação seletiva e/ou a supressão da publicação de resultados específicos, causando viés e consequentemente a redução da validade dos resultados da RS⁵¹.

Viés de publicação ocorre quando resultados de estudos elegíveis para uma RS não estão disponíveis, podendo ser devido à significância estatística (valor-p), à magnitude ou à direção dos resultados⁵². Existe evidência empírica de que a publicação de um estudo é influenciada pela significância estatística, magnitude e direção dos resultados⁵³. Estudos com resultados estatisticamente significantes ou positivos têm maior probabilidade de serem publicados (aproximadamente três a quatro vezes mais) do que estudos negativos^{54, 55}.

Diante disso, autores de RS devem fazer todos os esforços possíveis, por meio de busca extensa em diferentes bases de dados e literatura cinzenta, para incluir todos ou quase todos os estudos relevantes, sejam eles publicados em periódicos ou não, e assim minimizar o impacto do viés de publicação na RS. Por isso, é muito importante que, além das bases de dados tradicionais, seja realizada busca por registros de ensaios clínicos, resumos apresentados em congressos/conferências, relatos de agências regulatórias e relatos das empresas financiadoras⁵².

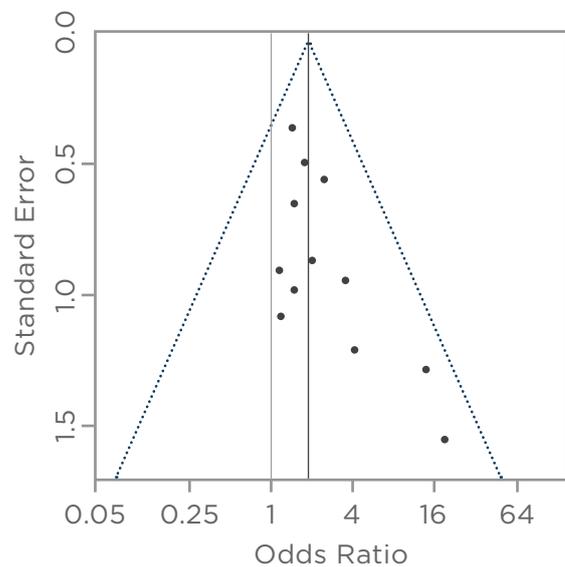
Em uma RS, foi encontrado que cerca de apenas 40% dos resumos apresentados em conferências foram publicados posteriormente por completo⁵⁶, sendo que os resumos com resultados estatisticamente favoráveis à intervenção experimental tinham maior probabilidade de serem publicados. Achados similares foram encontrados por Rising *et al.*⁵⁷, que avaliaram ensaios clínicos submetidos para a agência americana regulatória (FDA) e encontraram que ensaios clínicos com resultados favoráveis tinham maior probabilidade de serem publicados quando comparados com ensaios com resultados não favoráveis.

Além dos esforços para evitar e/ou minimizar a ocorrência do viés de publicação na RS, autores devem investigar a sua presença por meio da inspeção visual do gráfico do funil (do inglês, *funnel plot*) e/ou do gráfico do funil *contour-enhanced*, e dos testes estatísticos (Egger, Peters e Harbord)⁵⁸⁻⁶⁰ que avaliam assimetria do gráfico do funil (efeitos dos estudos pequenos). Para realizar a avaliação da simetria do gráfico do funil, são necessários alguns pré-requisitos: pelo menos dez estudos por desfecho; pelo menos um estudo com estimativa de efeito estatisticamente significativa; e inclusão de estudos com tamanhos de amostras diferentes⁵².

O gráfico do funil é um gráfico de dispersão das estimativas de efeito de cada estudo contra uma medida de precisão (erro-padrão é o mais indicado) ou tamanho da amostra (**Figura 3**)⁵². A precisão das estimativas de efeito tende a aumentar à medida que o tamanho da amostra aumenta. Assim, no gráfico do funil, as estimativas de efeito com menor precisão (estudos pequenos) irão normalmente se espalhar na base do gráfico, enquanto que as estimativas com maior precisão

(menores valores de erro-padrão) irão se afunilar no topo⁵². Quando o gráfico do funil é simétrico, os estudos com maior precisão (normalmente com maior amostra) ficam posicionados mais próximos do valor verdadeiro da estimativa de efeito, enquanto as estimativas dos estudos pequenos se espalham simetricamente de cada lado do gráfico. Quando há a suspeita de viés de publicação, porque estudos com amostras pequenas sem significância estatística não foram publicados, haverá uma assimetria do gráfico do funil, com ausência de estimativas de efeitos desses estudos em um lado da base do gráfico⁵². O gráfico do funil está disponível nas meta-análises realizadas no programa Review Manager, e também pode ser realizado por meio de comandos em outros programas, como R e Stata).

Figura 3 - Exemplo do gráfico do funil com linha do efeito nulo (linha cinza-claro) e linha do efeito sumário (linha cinza-escuro)

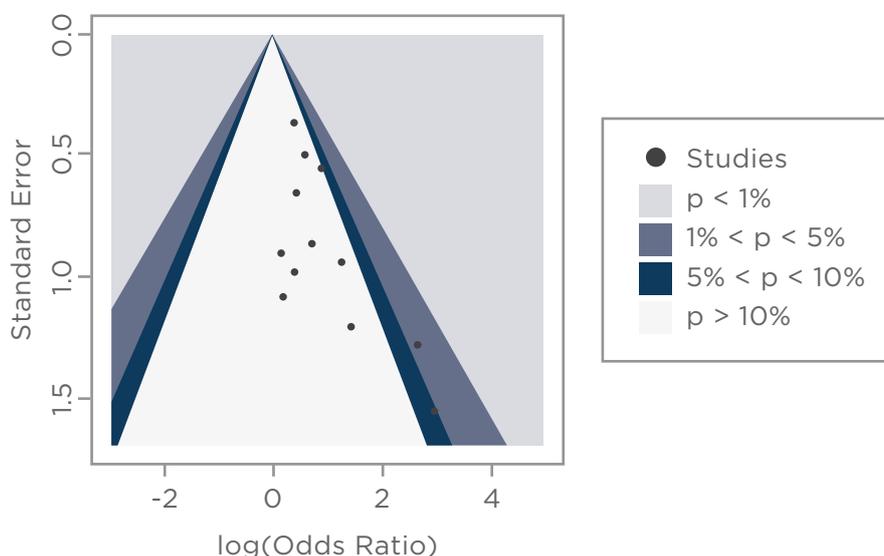


Fonte: Langan *et al.* (2012).
 Legenda: *Odds Ratio*: Razão de Chances;
Standard Error: Erro- Padrão.

Apesar de o gráfico do funil ser usado tradicionalmente para investigar a presença de viés de publicação em uma meta-análise, ele deve ser interpretado com cautela. Esse método possui baixo poder estatístico, dificuldade de interpretação por parte dos pesquisadores e diferentes explicações, além do viés de publicação, para a distribuição assimétrica dos estudos^{52, 61}. Essas outras razões que podem causar assimetria do gráfico do funil são: qualidade metodológica pobre, que leva a estimativas de efeito exageradas nos estudos pequenos; heterogeneidade entre os estudos; uso de métricas (**razão de chances** e diferença média padronizada, por exemplo) que estão naturalmente correlacionados com seus erros-padrão; e o acaso (chance)⁵².

Diante dessa dificuldade em atribuir a assimetria do gráfico do funil à presença de viés de publicação, foi incorporado a ele linhas de contorno correspondente à significância estatística ($p = 0,01, 0,05, 0,1$), o qual é chamado de gráfico do funil *contour-enhanced* (**Figura 4**)⁶². Esse método permite visualizar se as áreas em que há estudos faltantes são áreas de significância estatística das estimativas de efeitos ou não. Dessa forma, se há estudos faltantes em áreas em que os resultados seriam estatisticamente não significantes e desfavoráveis à intervenção experimental, pode-se atribuir isso à possibilidade da presença do viés de publicação. Na situação contrária, em que há estudos faltantes em áreas em que os resultados seriam estatisticamente significantes e favoráveis à intervenção experimental, há assimetria do gráfico, contudo é mais provável que esteja relacionada a outros fatores (previamente descritos no parágrafo acima) do que ao viés de publicação⁵².

Figura 4 - Exemplo do gráfico do funil com linhas de significância



Fonte: Langan *et al.* (2012).

Legenda: *Odds Ratio*: Razão de Chances; *Standard Error*: Erro- Padrão; *Studies*: Estudos.

Além da inspeção visual do gráfico do funil ou do gráfico do funil *contour-enhanced*, autores de RS também devem avaliar a assimetria do gráfico do funil usando testes estatísticos. Esses testes avaliam se a associação das estimativas de efeito e medidas de tamanho do estudo (precisão) são maiores do que esperado ao acaso, ou seja, se as estimativas dos estudos pequenos diferem sistematicamente dos estudos grandes. Entre os testes mais comumente utilizados estão: o teste de Egger (usado para desfechos contínuos)⁵⁸; o teste de Harbord (para desfechos dicotômicos em que o efeito de tratamento foi mensurado por **razão de chances** e risco relativo)⁵⁹; e o teste de Peters (apenas para **razão de chances**)⁶⁰. Os testes são dependentes do p-valor utilizado para definir a significância estatística, sendo recomendado utilizar um limiar de p-valor menor que 0,1. A hipótese nula desses testes indica que há simetria do gráfico do funil. Esses testes não são realizados no programa Review Manager, mas podem ser usados por meios de comandos simples nos programas R e Stata, entre outros.

Importante destacar que todos os testes estatísticos que avaliam viés de publicação possuem baixo poder estatístico e são dependentes do número de estudos incluídos em uma meta-análise. Esses testes tendem a subestimar a assimetria do gráfico do funil quando o número de estudos é pequeno, portanto não são confiáveis para descartar um possível viés de publicação nessas situações⁶³.

Assim, autores de RS devem sempre avaliar a possível presença de viés de publicação, utilizando a inspeção visual dos gráficos de funil (de preferência, usando o *contour-enhanced*) e os testes estatísticos para detectar a assimetria deles. Isso só deve ser realizado na presença de mais de dez estudos (e com tamanhos amostrais diferentes) na meta-análise, de modo que haja pelo

menos um estudo com significância estatística favorecendo a intervenção experimental⁵². Caso, após a avaliação, não seja constatada assimetria do gráfico do funil, pode haver certeza no conjunto da evidência. Se for detectada assimetria do gráfico do funil, o viés de publicação pode ser a causa e, dessa forma, há uma incerteza no conjunto da evidência (estimativas de efeitos podem estar superestimadas). Os autores devem realizar um processo de busca extensa por estudos não publicados, e as estimativas devem ser ajustadas para o viés de publicação⁵².

3.5.14 | Avaliação da qualidade da evidência

Qualidade da evidência diz respeito ao grau de confiança que se pode ter em uma determinada estimativa de efeito, ou seja, a confiança no resultado gerado pelo conjunto de evidências acerca de um determinado desfecho^{64, 65}. A avaliação da qualidade da evidência em uma RS é muito importante porque a utilidade de uma estimativa da magnitude do efeito de tratamento depende diretamente da confiança nessa estimativa. Dessa forma, esta atualização da diretriz de RS recomenda fortemente que toda RS realize a avaliação da qualidade usando o sistema GRADE, garantindo assim um processo sistemático e transparente nesse processo da RS.

O grupo GRADE desenvolveu um sistema para realizar a graduação da qualidade da evidência, o qual tem sido utilizado por diretrizes clínicas e RS. Esse sistema classifica as evidências em alta, moderada, baixa ou de muito baixa qualidade, levando em conta todos os fatores que determinam o quanto confiável são os resultados, como pode ser observado no **Quadro 9**^{64,65}. A classificação da evidência deve ser realizada para cada desfecho separadamente, já que provavelmente a qualidade da evidência irá diferir entre os desfechos⁶⁶.

Em RS de ECR, a qualidade da evidência sempre começa como “alta” (não há preocupações em nenhum dos fatores avaliados), podendo ser rebaixada em até três níveis (chegando à qualidade “muito baixa”), de acordo com a avaliação dos cinco domínios que compõem a avaliação pelo sistema GRADE^{64, 65}.

Quadro 9 – Avaliação da qualidade da evidência

Nível da evidência	Definição
Alta	Há muita confiança que o verdadeiro efeito está próximo daquele estimado.
Moderada	Há moderada confiança na estimativa de efeito: o verdadeiro efeito provavelmente está próximo daquele estimado, mas existe a possibilidade de ser substancialmente diferente.
Baixa	A confiança na estimativa de efeito é limitada: o verdadeiro efeito pode ser substancialmente diferente daquele estimado.
Muito baixa	Há muito pouca confiança na estimativa de efeito: o verdadeiro efeito provavelmente é substancialmente diferente daquele estimado.

Fonte: Elaboração GRADE *working group* - <http://www.gradeworkinggroup.org>.

3.5.14.1 Domínios que reduzem a qualidade da evidência

Os autores das RS devem descrever, com o máximo de transparência e detalhes, como será realizado o processo de avaliação da qualidade da evidência usando o sistema GRADE, assim como quantos autores participarão dessa etapa. O julgamento sobre os domínios que determinam a qualidade da evidência e a classificação da qualidade da evidência de cada desfecho devem ser descritos nos resultados, assim como relatados na tabela “Sumário dos Achados”.

Os domínios que podem reduzir a qualidade da evidência são: limitações metodológicas (risco de viés); inconsistência; evidência indireta; imprecisão; e viés de publicação. Para cada domínio, a qualidade da evidência pode ser rebaixada em um nível (se considerada um problema grave) ou dois níveis (se considerada um problema muito grave), conforme discutido a seguir. A **Figura 5** resume o processo de avaliação da qualidade da evidência.

Figura 5 – Passo a passo de como deve ser realizada a análise da qualidade de evidência em revisões sistemáticas de ensaios clínicos randomizados

Passo 1: defina o nível inicial da qualidade da evidência.

Delineamento do estudo	Certeza inicial no resultado
Ensaios clínicos randomizados	Alta certeza



Passo 2: considere reduzir a qualidade da evidência.

Razões para considerar diminuir a certeza no resultado	
Reduzir um nível (limitações sérias) ou dois níveis (limitações muito sérias)	Risco de viés Inconsistência Evidência indireta Imprecisão Viés de publicação



Passo 3: classificação final da qualidade da evidência.

Certeza na estimativa de efeito
Alta ⊕ ⊕ ⊕ ⊕
Moderada ⊕ ⊕ ⊕ ○
Baixa ⊕ ⊕ ○ ○
Muito baixa ⊕ ○ ○ ○

Fonte: Adaptado de Schunemann *et al.* (2019)⁶⁶.

Limitações metodológicas (risco de viés)

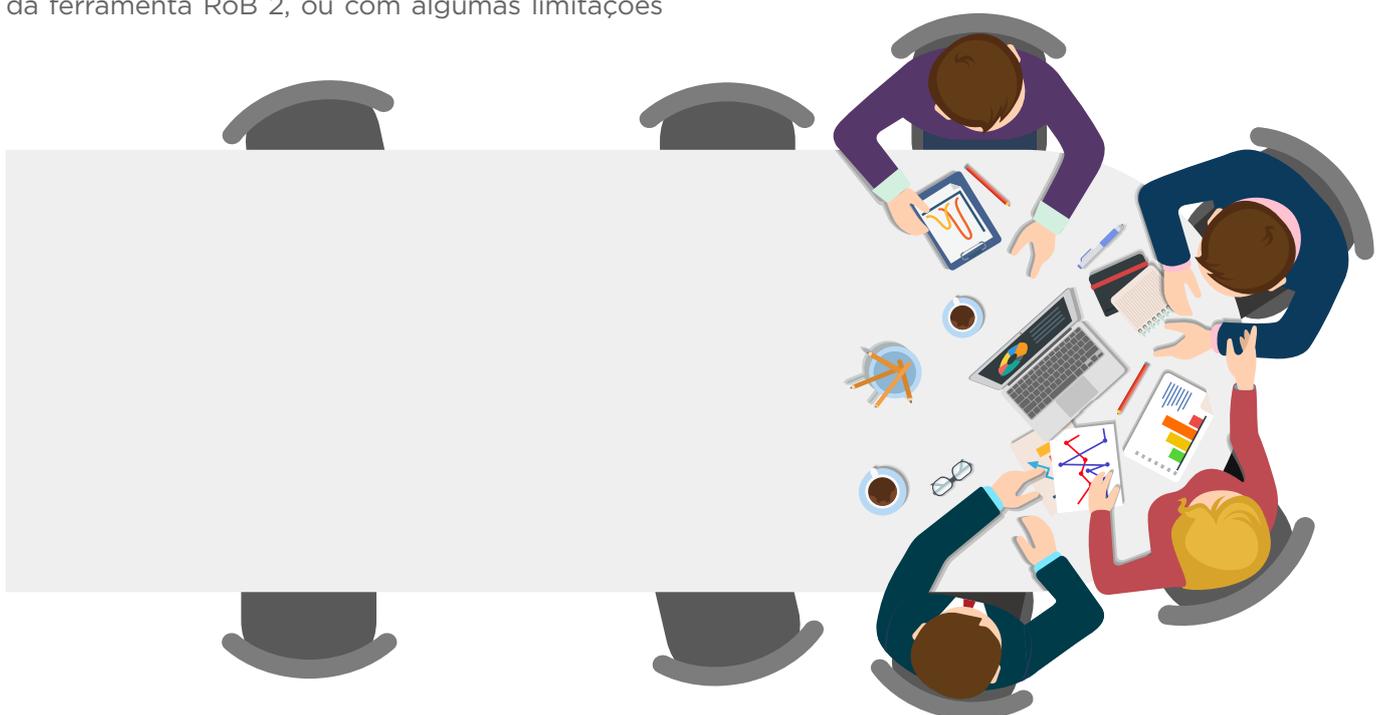
A qualidade da evidência pode ser reduzida em um ou até dois níveis, se houver limitações metodológicas importantes na maioria dos ensaios clínicos randomizados incluídos na meta-análise que provavelmente indiquem que o efeito da intervenção para determinado desfecho está sujeito a vieses⁶⁷. Essas limitações metodológicas são averiguadas durante a avaliação do risco de viés dos estudos incluídos pela ferramenta da Cochrane RoB 2 (conforme explicado no tópico anterior “Avaliação do risco de viés”).

Os autores da RS devem especificar o número de estudos ou a quantidade de informação fornecida para meta-análise que foram julgados como “alto risco de viés”, e também qual ou quais critérios (sequência de randomização, cegamento dos participantes etc.) foram levados em conta para reduzir a qualidade da evidência⁶⁶. A qualidade da evidência não será reduzida (ou seja, há alta confiança na evidência) quando a maioria dos estudos incluídos ou estudos com maior peso na meta-análise forem julgados como “baixo risco de viés”. A qualidade da evidência deve ser reduzida em um nível quando a maior parte da evidência é composta por estudos com limitação séria em um domínio da ferramenta RoB 2, ou com algumas limitações

para vários domínios. A qualidade da evidência pode ser reduzida em dois níveis quando a maior parte da evidência é composta por estudos com limitações muito sérias, ou seja, se foram julgados como “alto risco” para alguns domínios da ferramenta RoB 2⁶⁶.

Inconsistência (heterogeneidade não explicada)

Estudos podem obter estimativas de efeito de determinadas intervenções que diferem muito entre si, gerando o que se chama de heterogeneidade⁸. Por exemplo, os efeitos de tratamento podem mudar conforme a população (redução de mortalidade em pacientes com doença mais grave comparados com pacientes em fase inicial da doença). Em situações em que há boa evidência de que o efeito de tratamento dos desfechos pode ser diferente em análises de subgrupo, a síntese da evidência pode ser realizada em tabelas “Sumário dos Achados” separadas⁶⁶. Por exemplo, uma tabela para participantes com condição mais grave, e outra tabela para pacientes com condição mais leve. Quando a heterogeneidade existe, dificultando a interpretação dos resultados e não pode ser explicada, a qualidade da evidência deve ser diminuída⁶⁸.



Os autores da RS devem especificar se utilizaram a inspeção visual da sobreposição dos intervalos de confiança e/ou quais foram as medidas estatísticas (Chi^2 , I^2 , Tau^2) e os critérios adotados para julgar a heterogeneidade. Quando o I^2 for utilizado, os autores devem descrever a heterogeneidade como considerável, substancial, moderada ou não importante⁶⁶. Se os autores julgarem que não há inconsistência (heterogeneidade) importante, a qualidade da evidência não deve ser reduzida. Em caso de apenas um estudo incluído para avaliação do desfecho, os autores não devem reduzir a qualidade da evidência (não utilizar o termo “não aplicável”)⁶⁶. A qualidade da evidência deve ser reduzida em um nível quando há inconsistência importante, a qual não pode ser explicada, identificada pela ausência de sobreposição dos intervalos de confiança; teste do Chi^2 com p -valor < 0.1 ; e valor alto do I^2 ($> 50\%$ ou 60%). Se os autores julgarem que a inconsistência é muito importante (e não pode ser explicada), a qualidade da evidência deve ser reduzida em dois níveis⁶⁶.

Evidência indireta

A evidência é considerada indireta quando a questão abordada não é respondida diretamente pelos estudos disponíveis, seja por diferenças na população, nas intervenções, nos comparadores ou nos desfechos (PICO)⁶⁶. Existem situações em que a RS encontra ensaios clínicos randomizados que preenchem os critérios de inclusão, porém eles abordam apenas uma versão restrita da pergunta de pesquisa. Não se deve confundir esse ponto com as comparações indiretas que são realizadas nas RS com meta-análise em rede.

Para exemplificar, podemos imaginar uma RS que tem como objetivo avaliar a prevenção de obesidade em pacientes com osteoartrite do joelho, e a maioria dos estudos encontrados incluiu também participantes com osteoartrite do quadril. Nessa situação, a evidência seria re-

duzida em um nível, já que a questão principal é sobre osteoartrite do joelho apenas. Outro exemplo, agora relacionado à intervenção de interesse, seria uma RS que está interessada em avaliar se cirurgia de prótese de joelho é eficaz no tratamento da osteoartrite avançada do joelho, e os estudos encontrados incluíram apenas aqueles em que a cirurgia foi realizada por cirurgiões com mais de 20 anos de experiência. A evidência pode ser considerada indireta e a qualidade reduzida em um nível, já que não se aplica para todos os cirurgiões (com e sem experiência). A qualidade da evidência poderá ser reduzida um nível quando a evidência indireta for julgada como séria; e em dois níveis quando a evidência indireta for julgada como muito séria⁶⁹.

Imprecisão

Ocorre imprecisão e, conseqüentemente, diminuição na confiança do efeito estimado de uma intervenção em uma RS quando geralmente esta inclui estudos com poucos participantes, e/ou poucos eventos clínicos são observados, gerando intervalos de confiança largos que compreendem tanto benefício como dano ao paciente⁷⁰. Se a RS não atinge o critério de tamanho ótimo da informação (TOI) ou tamanho ótimo da revisão (TOR), ou seja, se o número total de participantes incluídos em uma RS é menor que o número de participantes sugerido pelo cálculo de tamanho amostral para um ensaio clínico com poder estatístico adequado, a qualidade da evidência deve ser reduzida⁶⁶. O cálculo do TOI ou TOR é realizado da mesma forma que se faz para o ensaio clínico. É necessário definir valores dos erros do tipo I (conhecido como α) e do tipo II (conhecido como β) (são recomendados 0,05 e 0,2, respectivamente), a diferença esperada entre os grupos e a taxa de eventos (para desfechos dicotômicos) ou o desvio-padrão da população (para desfechos contínuos)⁷⁰.

Quando os intervalos de confiança ao redor do efeito estimado são substancialmente largos, a qualidade da evidência também deve ser reduzida, mesmo que o TOI seja alcançado⁷⁰. Para desfechos contínuos, geralmente é considerada imprecisão quando o IC a 95% engloba tanto o efeito nulo como a diferença mínima clinicamente importante. Para desfechos dicotômicos, quando o IC a 95% compreende tanto o efeito nulo como considerável risco ou benefício (risco relativo $\geq 1,25$ ou $\leq 0,75$), a qualidade da evidência deve ser reduzida para imprecisão⁷⁰.

Quando há número pequeno de participantes ou poucos eventos, e o IC compreende tanto benefício como dano, a qualidade da evidência deve ser reduzida em dois níveis⁶⁶.

Viés de publicação

Quando ensaios clínicos não são publicados, e isso ocorre devido aos resultados deles, o viés de publicação pode estar presente⁷¹. Viés de publicação também é provável quando estudos com amostras pequenas apontam para benefícios acentuados, enquanto estudos com tamanho de amostra mais robusta demonstram pouco ou nenhum efeito da intervenção. Portanto RS baseada em estudos pequenos apresentam maior risco de viés de publicação. RS que incluem a maioria dos estudos com financiamento da indústria e amostra reduzida também levantam suspeitas sobre o viés de publicação⁶⁶. Nessas situações, a qualidade da evidência pode ser reduzida.

Os autores da RS devem indicar a razão e a forma utilizadas para avaliar a provável presença de viés de publicação, como o teste de assimetria do gráfico do funil ou pela utilização de testes estatísticos, como o teste de Egger, ou pela inclusão de estudos pequenos com resultados positivos⁶⁶. Explicação detalhada sobre avaliação do viés de publicação pode ser encontrada no tópico “Viés de publicação”.

A qualidade da evidência deve ser reduzida um nível quando há suspeita de viés de publicação, e em dois níveis quando há suspeita séria de viés de publicação. Quando o viés de publicação é improvável, como no caso de simetria do gráfico do funil e/ou a RS realizou uma busca extensa na literatura por estudos não publicados, a qualidade da evidência não deve ser reduzida⁶⁶.

3.5.14.2 Tabela “Sumário dos Achados”

A qualidade da evidência deve ser apresentada com a tabela “Sumário dos Achados”. Os autores necessitam predeterminar as comparações que serão apresentadas na tabela, priorizando os principais desfechos (primários e/ou de importância clínica para os participantes) da RS. É importante que cada tabela “Sumário dos Achados” contenha no máximo sete desfechos. Essa tabela é importante, porque ela resume as informações mais importantes dos achados principais da RS, como a magnitude dos efeitos relativos e absolutos da intervenção, e a evidência (robustez e qualidade) disponível para aquela comparação⁶⁶.



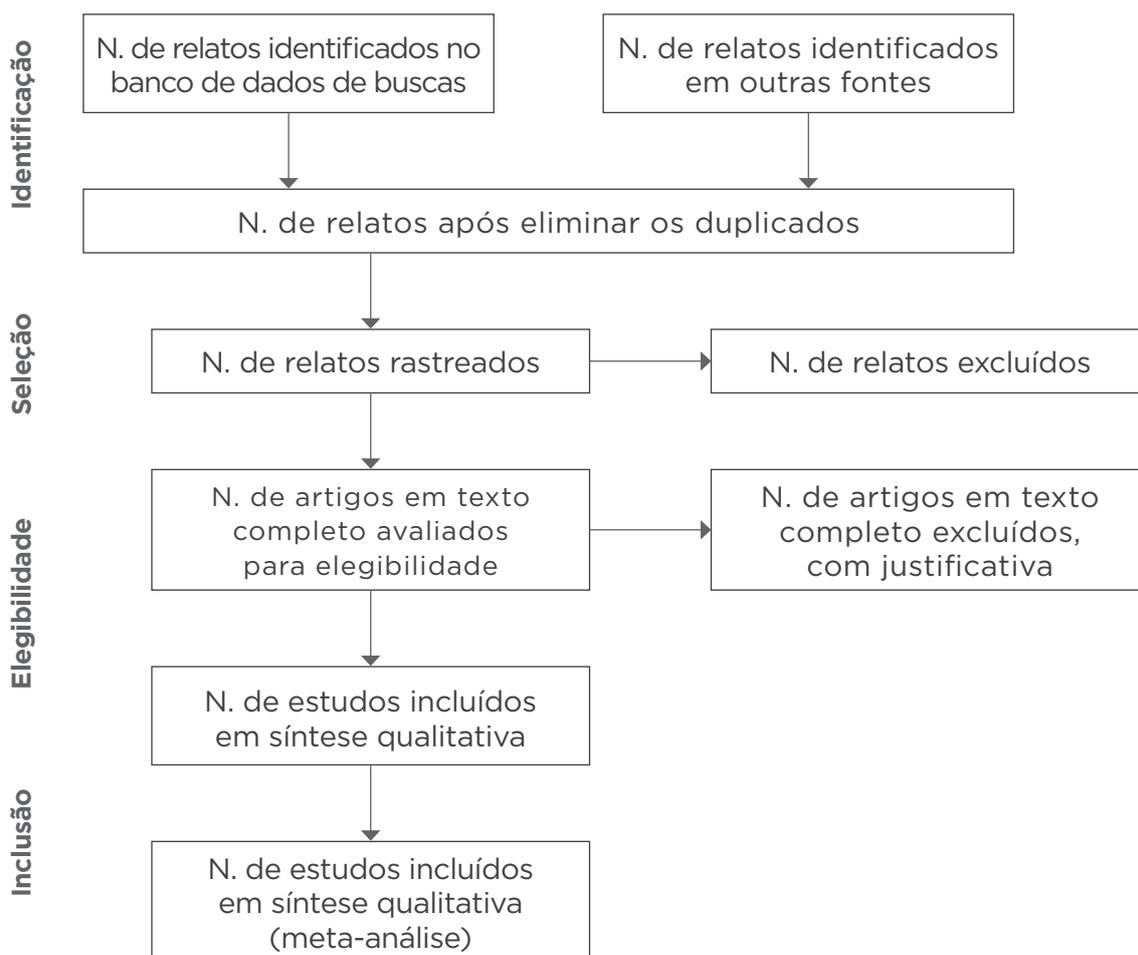
Os autores devem predeterminar as comparações que serão apresentadas na tabela.

3.6 | Resultados

3.6.1 | Estudos selecionados na revisão sistemática

Os autores da RS devem identificar quantos estudos foram elegíveis no processo de seleção (triagem de títulos e resumos e leitura de textos completos), quantos foram incluídos na RS, assim como identificar o número de estudos em andamento (registros de ensaios clínicos elegíveis). É primordial que os autores preencham o fluxograma de seleção, a fim de tornar o processo o mais transparente possível. Na **Figura 6**, encontra-se o fluxograma no modelo PRISMA, no qual os autores devem inserir o número de estudos encontrados na busca, o número de publicações após a remoção de duplicatas, e o número de estudos em cada fase de seleção, além das justificativas de exclusão na fase de leitura de textos completos. É recomendado que os autores apresentem em anexo uma lista dos estudos excluídos e uma de estudos em andamento (para aumentar a transparência do processo) após leitura completa com as devidas justificativas de exclusão.

Figura 6 - Fluxograma do processo de seleção dos estudos



Fonte: Tradução oficial do fluxograma PRISMA (<http://www.prisma-statement.org/Translations/Translations>).

3.6.2 | Caracterização dos estudos incluídos

É importante que os autores descrevam no texto, de forma resumida, as informações relevantes dos estudos incluídos, abordando cada componente do PICO. Além do relato no texto, recomenda-se que sejam apresentados, em formato de quadro, os dados extraídos com relação à caracterização dos estudos incluídos (**Figura 7**), informando autor; ano de publicação; país em que o estudo foi conduzido; cenário/contexto; informações referentes à população (tamanho da amostra, condição de saúde, idade, proporção de homens e mulheres e outras informações relevantes); características da intervenção e do comparador (dosagem/frequência/número de aplicações, de acordo com o tipo de tecnologia avaliada) – recomenda-se a utilização do TIDler checklist para auxiliar na descrição das intervenções (<https://www.equator-network.org/reporting-guidelines/tidier/>); fonte de financiamento; e declaração de conflito de interesse dos autores. Em outro quadro (**Figura 10**), os autores devem: identificar os desfechos avaliados nos estudos incluídos, sendo que estes devem ser os mesmos estabelecidos no PICO da RS; informar o tempo de seguimento em que o desfecho foi analisado; e fornecer a medida de efeito dos resultados dos desfechos fornecidos pelos estudos individuais. Os quadros podem ser adaptados de acordo com a pergunta de pesquisa e as características da população.

Os resultados devem ser baseados em todos os estudos elegíveis. No caso de os autores optarem por excluir algum estudo, essa decisão precisa ser justificada. Caso os autores tenham encontrado múltiplos estudos de um estudo principal, somente o estudo principal deve ser considerado, incluído e apresentado nos resultados. Diante disso, os autores podem identificar, no relato da RS, que as informações do estudo do(s) *Autor(es)* (*ano de publicação*) foram advindas de mais de uma publicação, identificando a referência delas no relato.

Figura 7 – Exemplo de quadro de caracterização dos ensaios clínicos randomizados incluídos

Autor, ano	Local	Número de indivíduos	População	Intervenção	Comparador	Desfechos primários e secundários	Financiamento
Autor, ano.	País onde o estudo foi conduzido; cenário e contexto.	Tamanho da amostra.	Condição de saúde, gênero, idade.	Nome da intervenção e característica (dosagem, frequência, número de aplicações, entre outros (de acordo com o tipo de intervenção).	Nome do comparador e características (dosagem, frequência, número de aplicações, entre outros (de acordo com o tipo de intervenção).	Desfechos primários. Desfechos secundários. Desfechos primários. Desfechos secundários.	

Fonte: Elaboração própria.

3.6.3 | Resultados da avaliação do risco de viés

Os resultados da avaliação do risco de viés dos estudos incluídos podem ser apresentados em formato de quadro, o qual inclui as respostas para cada pergunta sinalizadora dos domínios, o julgamento do risco de viés de cada domínio, com o suporte do julgamento para aumentar a transparência do processo de avaliação. Devido à extensão dessa tabela, recomenda-se que seja inserida como anexo do texto da RS. Recomenda-se também que o resumo da avaliação do risco de viés para cada estudo seja apresentado em formato de figura (**Figura 8**).

Figura 8 - Exemplo do resultado da avaliação do risco de viés usando a ferramenta RoB 2.0

	Risk of bias domains					Overall
	D1	D2	D3	D4	D5	
Estudo 1	+	+	+	+	+	+
Estudo 2	-	+	+	+	+	+
Estudo 3	?	+	-	+	+	-
Estudo 4	+	+	x	+	-	x
Estudo 5	x	x	+	+	-	+
Estudo 6	+	x	-	+	+	-
Estudo 7	+	-	-	x	+	-
Estudo 8	+	-	+	+	+	+
Estudo 9	+	+	x	+	+	x

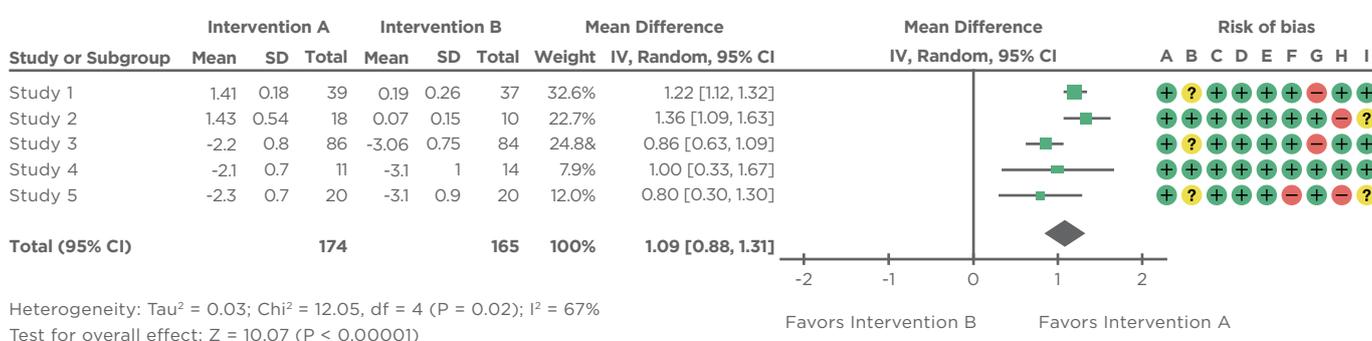
<p>Domains</p> <p>D1: Bias arising from the randomization process</p> <p>D2: Bias due to deviations from intended intervention</p> <p>D3: Bias due to missing outcome data</p> <p>D4: Bias in measurement of the outcome</p> <p>D5: Bias in selection of the reported result</p>	<p>Judgement</p> <p>x High</p> <p>- Some concerns</p> <p>+</p>
--	--

Fonte: Adaptado de McGuinness e Higgins (2020).

No texto, os autores da RS podem identificar quantos e quais estudos foram classificados como “baixo risco”, “alto risco” e “algumas preocupações”. Importante lembrar que os autores devem identificar para quais desfechos (de preferência, os desfechos primários ou de maior relevância para os pacientes) foram avaliados o risco de viés. Os autores também podem resumir as informações obtidas da avaliação do risco de viés para cada domínio (por exemplo, “sete^{1,2,5,8,11,12,13} de dez ECRs foram classificados como alto risco de viés para o domínio ‘processo de randomização’”).

No caso de revisões sistemática com meta-análise, é recomendado que os autores apresentem o julgamento do risco de viés no *forest plot* juntamente ao resultado de cada estudo incluído na meta-análise (**Figura 9**). Isso proporciona uma impressão visual da contribuição dos estudos com diferentes níveis de risco de viés, auxiliando na interpretação dos dados para o desfecho³⁶.

Figura 9 – Exemplo de *forest plot* com adição do resultado da avaliação do risco de viés



Risk of a Bias legend

- (A) Random sequence generation (selection bias)
- (B) Allocation concealment (selection bias)
- (C) Baseline imbalances (selection bias)
- (D) Blinding of participants (performance bias)
- (E) Blinding of personnel (performance bias)
- (F) Blinding of outcome assessment (detection bias)
- (G) ITT analysis (attrition bias)
- (H) Incomplete outcome data (attrition bias)
- (I) Selective reporting (reporting bias)

Fonte: Elaboração própria.

3.6.4 | Resultados dos estudos individuais e/ou das meta-análises

Os resultados das RS devem ser apresentados de maneira ordenada e sistematizada. É necessário que os autores estruturarem o texto de acordo com a ordem da importância dos desfechos incluídos em cada comparação, ou seja, cada desfecho deverá ter os seus resultados apresentados para todos os estudos que apresentem dados sobre ele. Os autores devem apresentar primeiramente os resultados dos desfechos primários, seguido então pelos resultados dos desfechos secundários. Caso não sejam encontrados estudos que avaliaram algum desfecho ou alguma comparação, essa informação deve ser relatada no texto. Autores de RS devem sempre apresentar resultados de cada estudo incluído e/ou meta-análise de todos os desfechos e comparações especificadas no PICO.

3.6.4.1 Estudos individuais

Quando meta-análises não forem realizadas, os autores precisam justificar por que isso não foi possível. Os resultados devem ser apresentados por meio de uma tabela de associação entre os estudos e os desfechos avaliados, que deve conter os resultados de cada grupo de intervenção, assim como o efeito do tratamento (diferença entre os grupos) com o IC (**Figura 10**). Os autores também podem optar por apresentar os resultados individuais em um *forest plot* (o qual, nessa situação, não conterá o efeito de tratamento agrupado), o que facilita a interpretação visual por parte dos leitores⁵. No texto, os autores podem descrever os resultados indicando se são favoráveis a algum grupo de intervenção ou não, assim como a classificação em relação à qualidade da evidência e o tamanho amostral (ex.: “Existe baixa qualidade da evidência (reduzida um nível devido ao risco de viés e um nível devido à imprecisão) proveniente de um estudo, sugerindo que há um efeito clinicamente importante do grupo tratado com exercício na redução da dor em curto prazo comparado com o grupo que recebeu intervenção mínima. A dor foi de 5.5 pontos no grupo tratado com intervenção mínima (em uma escala de 0 a 10) e 3.5 pontos menor no grupo tratado com exercício (IC 95%, -1.1 a -5.3; n = 80)”).

É importante que os autores da RS apresentem os resultados de cada grupo de tratamento, (e não apenas o efeito de tratamento com o IC), já que leitores podem estar interessados, por exemplo, na variação da taxa de eventos entre os estudos, não apenas na estimativa do RR ou OR⁵. Além disso, essa apresentação permite a identificação de possíveis erros (que não deveriam acontecer, mas são relativamente comuns) no processo de extração dos dados da RS. Normalmente, os dados de cada grupo de intervenção apresentados são: número de eventos e tamanho amostral para dados dicotômicos; média, desvio-padrão e tamanho amostral para dados contínuos. Se os autores realizaram imputação de dados, como para a ausência de desvio-padrão de algum estudo, isso deve estar identificado no texto em que foi efetuada a imputação⁵.

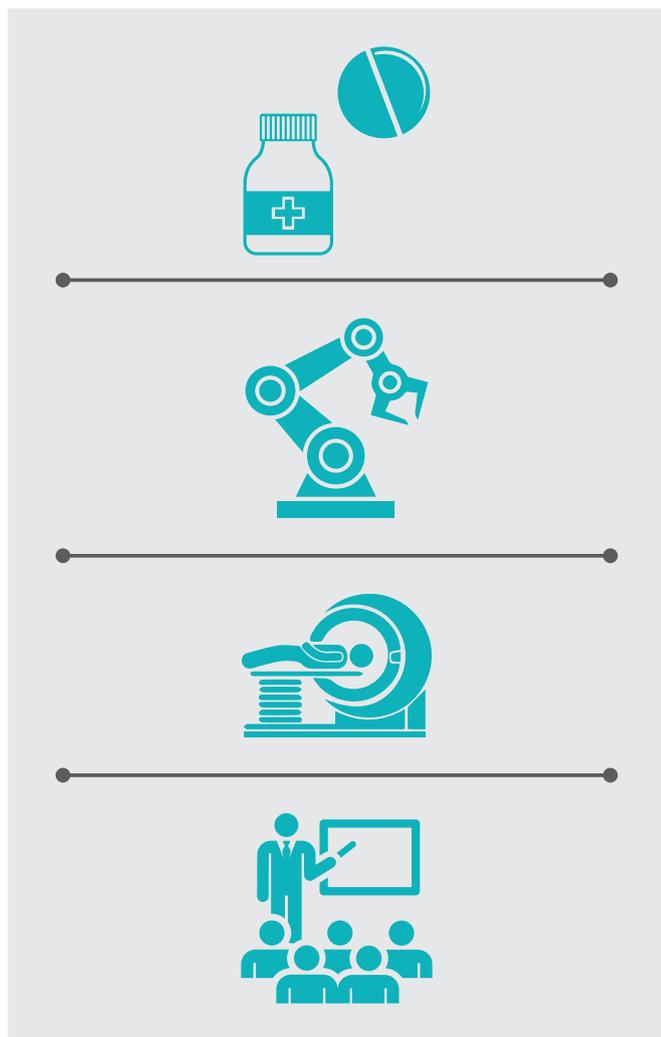
Figura 10 – Exemplo de quadro de resultados dos desfechos dos estudos incluídos por comparação

Intervenção A versus Intervenção B				
Estudos	Desfecho 1	Desfecho 2	Desfecho 3	Desfecho 4
Estudo 1	Número de eventos (número total de indivíduos) RR (IC 95%)			
	Média (desvio-padrão) DM (IC 95%)			
Estudo 2	Número de eventos (número total de indivíduos) RR (IC 95%)			
	Média (desvio-padrão) DM (IC 95%)			

Fonte: Elaboração própria.

3.6.4.2 Meta-análises

Quando meta-análises foram realizadas, os autores devem apresentar, com o efeito de tratamento agrupado e o IC, uma medida de inconsistência (I^2 , por exemplo), número de estudos e participantes incluídos em cada análise⁵. As meta-análises devem ser apresentadas em *forest plots*, os quais devem conter todas as informações mencionadas, além de demonstrar a contribuição (peso) de cada estudo para meta-análise. Os resultados apresentados em *forest plots* permitem identificar possível variação dos resultados dos estudos individuais incluídos, além de facilitar a interpretação dos resultados por parte dos leitores, inclusive da variação dos resultados individuais incluídos.



No texto, os autores da RS devem descrever os resultados, indicando se são favoráveis a algum grupo de intervenção ou não, além das informações citadas anteriormente. Um exemplo que demonstra como o resultado deve ser relatado pode ser visto a seguir: “Existe moderada qualidade de evidência (reduzida um nível devido à inconsistência) de que o grupo tratado com cirurgia provavelmente reduz a incapacidade no seguimento de um ano comparado com o grupo placebo. A incapacidade foi 53.9 pontos (em uma escala de 0 a 100), no grupo placebo, e 19 pontos menor no grupo tratado com cirurgia (IC 95%, -30.3 a -7.8; $I^2 = 50\%$, 3 estudos, n = 290 participantes).

3.6.4.3 Resultados das análises adicionais (subgrupo, sensibilidade e metarregressão)

Os autores devem relatar os resultados das análises de subgrupo, metarregressão e sensibilidade, seguindo a mesma sequência lógica dos resultados das análises principais. Os autores também devem informar se as análises adicionais foram pré-especificadas ou não. Para os resultados das análises de subgrupo, é necessário que os autores relatem o resultado do teste de interação entre os grupos, além dos dados de efeito de tratamento e IC de cada subgrupo⁵. Para os resultados das análises de metarregressão, os autores devem relatar o coeficiente de regressão (tamanho do efeito de tratamento) com o IC (e não apenas o p-valor)⁵. Já para os resultados da análise de sensibilidade, é necessário que os autores relatem também o efeito da intervenção com o IC. Eles devem, ainda, informar se os resultados são similares aos da análise principal ou não, para que seja possível avaliar a robustez deles.

3.6.5 | Resultado da avaliação da qualidade da evidência e tabela “Sumário dos Achados”

O resultado da avaliação da qualidade da evidência por meio do sistema GRADE deve ser apresentado em forma de um quadro, denominado quadro “Sumário dos Achados”, para cada comparação de interesse da RS. É recomendado que esse quadro da comparação principal da revisão apareça antes da introdução, já que se trata do sumário do principal achado da revisão. Caso haja outros quadros, eles devem aparecer na seção de resultados. Como já apresentado no tópico anterior, no texto, os autores devem sempre relatar os resultados das meta-análises ou estudos individuais (caso não seja possível a realização de uma meta-análise), acompanhados da qualidade da evidência daquele achado e para quais domínios houve redução da qualidade (caso haja redução).

O quadro “Sumário dos Achados” deve conter: 1) breve descrição da população e grupos de tratamento avaliados; 2) os principais desfechos de interesse dos pacientes (eficácia e segurança, sendo recomendados até sete desfechos); 3) uma medida do impacto de cada desfecho no grupo comparador; 4) magnitudes do efeito relativo e absoluto com o IC de 95%; 5) número de estudos incluídos e número de participantes para cada desfecho; 6) classificação da qualidade da evidência (alta, moderada, baixa ou muito baixa) para cada desfecho e indicação de para quais domínios a qualidade da evidência foi rebaixada ou aumentada, com as respectivas justificativas (**Figura 11**)⁶⁶. O quadro “Sumário dos Achados” pode ser feito no software gratuito GRADEpro (<https://gradepr.org/>).

Figura 11 – Exemplo de quadro-sumário dos resultados para avaliação da qualidade da evidência

Avaliação da qualidade da evidência (GRADE)							Resumo dos achados				Qualidade da evidência
Estudos	Delineamento	Limitações metodológicas	Inconsistência	Evidência indireta	Imprecisão	Viés de publicação	N. de participantes		Efeito		
							Tecnologia avaliada	Comparador	Relativo (IC 95%)	Absoluto	
Desfecho A (tempo de seguimento)											
2	ECR	Sem limitações graves	Sem inconsistência grave	Sem evidência indireta	Grave ¹	Não avaliado					Moderada
Desfecho B (tempo de seguimento)											
4	ECR	Grave ²	Sem inconsistência grave	Sem evidência indireta	Grave ³	Não avaliado					Baixa

¹Justificar a decisão.

²Justificar a decisão.

³Justificar a decisão.

Fonte: Adaptado do GRADEpro e das *Diretrizes Metodológicas: Sistema GRADE – manual de graduação da qualidade da evidência e força de recomendação para tomada de decisão em saúde*⁷².

3.6.6 | Interpretação dos resultados

É muito importante que os autores de uma RS interpretem adequadamente os resultados encontrados. Uma interpretação equivocada pode passar informações errôneas para os tomadores de decisão. Conseqüentemente, intervenções podem passar a ser utilizadas ou não, com base nessas informações.

Os resultados de uma RS, seja com realização de meta-análises ou não (resultados de estudos individuais), sempre devem ser relatados com uma estimativa pontual do efeito e o IC. A estimativa pontual (RR, OR, DM, DMP etc.) representa a melhor estimativa da magnitude e direção do efeito de uma intervenção, em comparação com outra intervenção (ou nenhum tratamento)⁷³. Já o IC representa a incerteza da estimativa pontual encontrada, e pode ser interpretado, de uma forma simples, como o intervalo de valores que possibilitará ter alguma certeza de que o verdadeiro valor se encontra⁷³. Caso o IC seja estreito, pode-se dizer que o efeito de tratamento tem certa precisão. Se o IC for muito largo, pode-se assumir que pouco se sabe sobre o efeito da intervenção, já que a imprecisão é muito grande; conseqüentemente, a incerteza sobre a evidência também⁷³.



Os autores da RS devem considerar a importância clínica dos achados no momento de interpretar a magnitude de efeito da intervenção e a precisão do IC⁷³. Para os autores afirmarem se uma intervenção é útil ou não, é preciso contextualizar clinicamente a situação. Por exemplo, suponha-se que uma RS esteja avaliando se uma intervenção reduz o nível de dor, e os autores relatam que uma diferença clinicamente importante entre os grupos seria de 2 pontos (em uma escala de 0 a 10). A meta-análise gerou uma DM de -3 pontos (estimativa pontual) e um IC a 95% de -2.2 a -3.5. Nessa situação, os autores poderiam afirmar que a intervenção foi útil em reduzir a dor, já que tanto a estimativa pontual e o IC extrapolam o critério para ser considerado clinicamente importante. Se a meta-análise tivesse gerado a mesma estimativa pontual (DM -3 pontos), mas com um IC mais largo (IC 95%, -1.5 a -4.5), os autores poderiam ainda concluir que a intervenção foi útil em reduzir a dor, mas com certo grau de imprecisão, já que o IC inclui o limiar considerado clinicamente importante; ou seja, o efeito do tratamento é compatível com a redução da dor entre -1.5 a -4.5 pontos. Se o IC for ainda mais largo, incluindo o efeito nulo da diferença igual a zero, os autores podem afirmar que existe a possibilidade de a intervenção não ter nenhum efeito e serem bastante céticos em relação à utilidade dela⁷³.

Quando os autores da RS optarem por utilizada a medida de efeito DMP para os desfechos contínuos, precisam definir a priori, no protocolo, como irão interpretar seus achados, já que normalmente tomadores de decisão não têm muita familiaridade com esse tipo de medida. Existem duas opções mais comumente utilizadas. Uma delas é apresentar diretamente os resultados em unidades de desvio-padrão e utilizar as regras de interpretação sugerida de Cohen e suas variações, que definem, por exemplo, a magnitude do efeito em peque-

no (< 0.40), moderado (0.40 a 0.70) e grande (> 0.70)⁷⁴. Porém essa interpretação é arbitrária (não é baseada em evidência empírica) e não leva em conta o contexto da população e a intervenção⁷³. Outra opção para interpretar uma medida de DMP é reexpressar essa medida em DM de um instrumento de mensuração utilizado nos estudos incluídos. Isso é realizado multiplicando-se o desvio-padrão do grupo-controle (normalmente, o estudo mais representativo e com menor risco de viés) pela DMP encontrada⁷³. Essa opção de transformar a DMP em unidades de medidas mais conhecidas pode ser mais útil para os tomadores de decisão.

Não é recomendado que autores de RS descrevam seus resultados como estatisticamente significantes ou não. Caso os autores optem por relatar o valor-p em seus resultados, eles devem relatar o IC com o valor-p exato encontrado⁷³. É comum que o valor-p encontrado seja interpretado inadequadamente. O valor-p é a probabilidade de obter um efeito igual ou maior que o observado, assumindo que a hipótese nula de não efeito seja verdadeira⁷³. É importante lembrar que o valor-p identifica o quão provável seria encontrar uma diferença ao acaso assumindo que esta não exista, portanto não fornece informações diretas sobre a magnitude do efeito de tratamento de uma intervenção em relação à outra.

3.7 | Discussão

3.7.1 | Resumo dos principais achados da revisão

Os autores da RS devem iniciar a discussão apresentando um resumo dos principais achados da revisão (evitando repetição detalhada do que foi descrito nos resultados). É recomendado que os autores interpretem os resultados indicando a relevância clínica e a qualidade da evidência para os principais desfechos de eficácia e de segurança. Os autores devem abordar, também, a aplicabilidade dos achados, se podem ser extrapolados para um contexto específico ou um contexto geral, e como a revisão poderá contribuir para a tomada de decisão por parte dos clínicos e gestores de saúde^{5, 75}.

3.7.2 | Limitações e pontos fortes

Na discussão da RS, os autores devem apresentar os pontos fortes realizados, como a busca abrangente em diversas bases de dados (incluindo bases de registros de ECR), a realização adequada de todos os processos (como avaliação do risco de viés, avaliação da qualidade da evidência), entre outros pontos. Além disso, é imprescindível que os autores identifiquem as possíveis limitações da RS e discutam como essas limitações podem ter influenciado os resultados encontrados e as possíveis implicações para a tomada de decisão e realização de futuras pesquisas. As limitações de uma RS podem surgir de algum processo da revisão, como a restrição de idioma na seleção dos estudos, ou podem estar relacionadas à generalização dos achados da revisão. As limitações desses achados

também podem estar relacionadas à quantidade e à qualidade (presença de vieses) dos estudos incluídos^{5, 75}.

3.7.3 | Comparações com outros estudos na literatura

Outro aspecto que deve ser abordado na discussão de uma RS é o quanto os resultados corroboram ou se contrapõem a achados de RS (e outros delineamentos, caso os autores acreditem ser pertinentes) previamente publicadas. Os autores devem discutir sucintamente similaridades e diferenças entre os estudos que possam explicar tais achados, como diferenças nos critérios de elegibilidade, forma que foram realizadas as buscas e como os dados foram sumarizados^{5, 75}.



3.8 | Conclusão

3.8.1 | Implicações para a pesquisa

Os autores de RS também devem orientar decisões sobre futuras pesquisas a serem realizadas sobre a intervenção e a condição pesquisada. Devem ser discutidos aspectos metodológicos, como de que modo os futuros ensaios clínicos devem ser realizados, a necessidade de amostras mais robustas, melhor descrição das intervenções ou um tempo de seguimento maior dos pacientes. Essas sugestões podem vir da avaliação do risco de viés e de fatores que reduziram a qualidade de evidência (GRADE) realizada pela RS^{5, 75}.

É importante que os autores também realizem um panorama dos estudos em andamento identificados para ajudar na decisão de orientar a realização de novos estudos ou não. Expressões gerais como “futuras pesquisas devem ser mais bem conduzidas” ou “mais pesquisas são necessárias” devem ser evitadas⁷⁵.

3.8.2 | Implicações para a prática

Nessa parte da RS, é necessário que os autores realizem uma interpretação geral da evidência, com informações para ajudar na tomada de decisão clínica e política. As implicações para a prática devem estar no texto de forma prática e clara, sempre restrito aos resultados encontrados na revisão. Os autores da RS também devem discutir possíveis fatores que possam auxiliar na tomada de decisão, como os valores relativos dos prováveis benefícios e riscos da intervenção, participantes em diferentes níveis de risco e aspectos relacionados aos recursos necessários para o uso da intervenção^{5, 75}. É importante que os autores da revisão sejam sempre cautelosos ao tirar conclusões sobre as implicações do estudo e não façam recomendações diretas sobre o uso ou não das intervenções⁷³.

Referências

1. COOK, D. J.; MULROW, C. D.; HAYNES, R. B. Systematic reviews: synthesis of best evidence for clinical decisions. **Annals of internal medicine**, v. 126, n. 5, p. 376-380, 1997.
2. LASSERSON, T. J.; THOMAS, J.; HIGGINS, J. P. T. Chapter 1: Starting a review. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
3. CHANDLER, J. *et al.* Chapter 1: Introduction. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated August 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
4. AROMATARIS, E.; MUNN, Z. (ed.). **Joanna Briggs Institute Reviewer's Manual**. Adelaide: The Joanna Briggs Institute, 2017. Disponível em: <https://reviewersmanual.joannabriggs.org/>. Acesso em: 12 maio 2020.
5. LIBERATI, A. *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. **BMJ**, v. 339, p. b2700, 2009.
6. HIGGINS, J. P. T. *et al.* **Methodological Expectations of Cochrane Intervention Reviews**. London: Cochrane, Version Oct. 2019.
7. CENTRE FOR REVIEWS AND DISSEMINATION. **Systematic Reviews: CRD's guidance for undertaking reviews in health care** [Internet]. 2020. Disponível em: www.york.ac.uk/inst/crd. Acesso em: 12 maio 2020.
8. BORENSTEIN, Michael *et al.* **Introduction to Meta-Analysis**. Hoboken (NJ): John Wiley & Sons, Ltd., 2009.
9. EGGER, M.; SMITH, G. D. Meta-Analysis: potentials and promise. **BMJ**, v. 315, n. 7119, p. 1371-1374, 1997.
10. O'ROURKE, K.; DETSKY, A. S. Meta-analysis in medical research: strong encouragement for higher quality in individual research efforts. **Journal of clinical epidemiology**, v. 42, n. 10, p. 1021-1024, 1989.
11. CHALMERS, I. *et al.* How to increase value and reduce waste when research priorities are set. **Lancet**, v. 383, n. 9912, p. 156-165, 2014.
12. POLLOCK, A. *et al.* Selecting and implementing overview methods: implications from five exemplar overviews. **Systematic reviews**, v. 6, n. 1, p. 145, 2017.

13. GARNER, P. *et al.* When and how to update systematic reviews: consensus and checklist. **BMJ**, v. 354, p. i3507, 2016.
14. SHAMSEER, L. *et al.* Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. **BMJ**, v. 350, p. g7647, 2015.
15. STEWART, L.; MOHER, D.; SHEKELLE, P. Why prospective registration of systematic reviews makes sense. **Systematic reviews**, v. 1, p. 7, 2012.
16. KIRKHAM, J. J.; ALTMAN, D. G.; WILLIAMSON, P. R. Bias due to changes in specified outcomes during the systematic review process. **PloS one**, v. 5, n. 3, p. e9810, 2010.
17. TRICCO, A. C. *et al.* A third of systematic reviews changed or did not specify the primary outcome: a PROSPERO register study. **Journal of clinical epidemiology**, v. 79, p. 46-54, 2016.
18. RUNJIC, E. *et al.* Half of systematic reviews about pain registered in PROSPERO were not published and the majority had inaccurate status. **Journal of clinical epidemiology**, v. 116, p. 114-121, 2019.
19. MOHER, D. *et al.* Helping editors, peer reviewers and authors improve the clarity, completeness and transparency of reporting health research. **BMC medicine**, v. 6, p. 13, 2008.
20. WEN, J. *et al.* The reporting quality of meta-analyses improves: a random sampling study. **Journal of clinical epidemiology**, v. 61, n. 8, p. 770-775, 2008.
21. SHEA, B. J. *et al.* AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. **BMJ**, v. 358, p. j4008, 2017.
22. WHITING, P. *et al.* ROBIS: A new tool to assess risk of bias in systematic reviews was developed. **Journal of clinical epidemiology**, v. 69, p. 225-234, 2016.
23. HIGGINS, J. P. T. *et al.* Chapter 8: Assessing risk of bias in a randomized trial. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
24. MCKENZIE, J. E. *et al.* Chapter 3: Defining the criteria for including studies and how they will be grouped for the synthesis. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.

25. PAEZ, A. Grey literature: An important resource in systematic reviews. **Journal of evidence-based medicine**, v. 10, n. 3, p. 233-240, Aug. 2017.
26. MCGOWAN, J. *et al.* PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. **Journal of clinical epidemiology**, v. 75, p. 40-46, 2016.
27. LEFEBVRE, C. *et al.* Chapter 4: Searching for and selecting studies. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
28. MARSHALL, I. J. *et al.* Rapid reviews may produce different results to systematic reviews: a meta-epidemiological study. **Journal of clinical epidemiology**, v. 109, p. 30-41, 2019.
29. TOLSGAARD, M. G. *et al.* Salami-slicing and plagiarism: How should we respond? **Advances in health sciences education : theory and practice**, v. 24, n. 1, p. 3-14, 2019.
30. MATHES, T.; KLASSEN, P.; PIEPER, D. Frequency of data extraction errors and methods to increase data extraction quality: a methodological review. **BMC Med. Res. Methodol.**, v. 17, n. 1, p. 152, 2017.
31. ELAMIN, M. B. *et al.* Choice of data extraction tools for systematic reviews depends on resources and review complexity. **Journal of clinical epidemiology**, v. 62, n. 5, p. 506-510, 2009.
32. LI, T.; HIGGINS, J. P. T.; DEEKS, J. J. (ed.). Chapter 5: Collecting data. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
33. JELICIC KADIC, A. *et al.* Extracting data from figures with software was faster, with higher interrater reliability than manual extraction. **Journal of clinical epidemiology**, v. 74, p. 119-123, 2016.
34. HIGGINS, J. P. T.; LI, T.; DEEKS, J. J. Chapter 6: Choosing effect measures and computing estimates of effect. *In*: HIGGINS, J. P. T. (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
35. WAN, X. *et al.* Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. **BMC medical research methodology**, v. 14, p. 135, 2014.
36. BOUTRON, I. *et al.* Chapter 7: Considering bias and conflicts of interest among the included studies. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.

37. SAVOVIC, J. *et al.* Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. **Annals of internal medicine**, v. 157, n. 6, p. 429-438, 2012.
38. JUNI, P.; ALTMAN, D. G.; EGGER, M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. **BMJ**, v. 323, n. 7303, p. 42-46, 2001.
39. STERNE, J. A. C. *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. **BMJ**, v. 366, p. l4898, 2019.
40. HERNAN, M. A.; ROBINS, J. M. Per-Protocol Analyses of Pragmatic Trials. **The New England journal of medicine**, v. 377, n. 14, p. 1391-1398. , 2017
41. VALENTINE, J. C. *et al.* Synthesizing Evidence in Public Policy Contexts: The Challenge of Synthesis When There Are Only a Few Studies. **Evaluation review**, v. 41, n. 1, p. 3-26, 2017.
42. CAMPBELL, M. *et al.* Lack of transparency in reporting narrative synthesis of quantitative data: a methodological assessment of systematic reviews. **Journal of clinical epidemiology**, v. 105, p. 1-9, 2019.
43. LOUGUIN, T. M. A systematic comparison of methods for combining p-values from independent tests. **Comp Stat Data Anal**, v. 47, n. 3, p. 467-485, 2004.
44. HARRISON S. *et al.* The albatross plot: A novel graphical tool for presenting results of diversely reported studies in a systematic review. **Research synthesis methods**, v. 8, n. 3, p. 281-289, 2017.
45. MCKENZIE, J. E.; BRENNAN, S. E. Chapter 12: Synthesizing and presenting findings using other methods. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
46. OGILVIE, D. *et al.* The harvest plot: a method for synthesising evidence about the differential effects of interventions. **BMC Med Res Methodol.**, v. 8, p. 8, 2008.
47. DEEKS, J. J.; HIGGINS, J. P. T.; ALTMAN, D. G. (ed.). Chapter 10: Analysing data and undertaking meta-analyses. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
48. 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 ensaios clínicos randomizados**. Brasília, DF: Ministério da Saúde, 2012.

49. FLETCHER, J. What is heterogeneity and is it important? **BMJ**, v. 334, n. 7584, p. 94-96, 2007.
50. RICHARDSON, M.; GARNER, P.; DONEGAN, S. Interpretation of subgroup analyses in systematic reviews: A tutorial. **Clin Epidemiol Global Health**, v. 7, n. 2, p. 192-198, 2019.
51. DEEKS, J. J.; MACASKILL, P.; IRWIG, L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. **Journal of clinical epidemiology**, v. 58, n. 9, p. 882-893, 2005.
52. PAGE, M. J.; HIGGINS, J. P. T.; STERNE, J. A. C. Chapter 13: Assessing risk of bias due to missing results in a synthesis. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
53. CHAN, A. W. *et al.* Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. **Jama**, v. 291, n. 20, p. 2457-2465, 2004.
54. HOPEWELL, S. *et al.* Time to publication for results of clinical trials. **The Cochrane database of systematic reviews**, n. 2, p. MR000011, 2007.
55. SCHMUCKER, C. *et al.* Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. **PLoS one**, v. 9, n. 12, p. e114023, 2014.
56. SCHERER, R. W. *et al.* Full publication of results initially presented in abstracts. **The Cochrane database of systematic reviews**, v. 11, p. MR000005, 2018.
57. RISING, K.; BACCHETTI, P.; BERO, L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. **PLoS medicine**, v. 5, n. 11, p. e217, 2008. Discussion e.
58. EGGER, M. *et al.* Bias in meta-analysis detected by a simple, graphical test. **BMJ**, v. 315, n. 7109, p. 629-6234, 1997.
59. HARBORD, R. M.; EGGER, M.; STERNE, J. A. A modified test for small-study effects in meta-analyses of controlled trials with binary endpoints. **Statistics in medicine**, v. 25, n. 20, p. 344334-57, 2006.
60. PETERS, J. L. *et al.* Comparison of two methods to detect publication bias in meta-analysis. **Jama**, v. 295, n. 6, p. 676-680, 2006.
61. TERRIN, N.; SCHMID, C. H.; LAU, J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. **Journal of clinical epidemiology**, v. 58, n. 9, p. 894-901, 2005.

62. PETERS, J. L. *et al.* Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. **Journal of clinical epidemiology**, v. 61, n. 10, p. 991-996, 2008.
63. FURUYA-KANAMORI, L. *et al.* P value-driven methods were underpowered to detect publication bias: analysis of Cochrane review meta-analyses. **Journal of clinical epidemiology**, v. 118, p. 86-92, 2020.
64. GUYATT, G. H. *et al.* GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. **BMJ**, v. 336, n. 7650, p. 924-926, 2008.
65. GRADE WORKING GROUP. **What is GRADE?** @2020. Disponível em: <https://www.gradeworkinggroup.org/>. Acesso em 15 jan. 2020.
66. SCHÜNEMANN, H. J. *et al.* Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.
67. GUYATT, G. H. *et al.* GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). **Journal of clinical epidemiology**, v. 64, n. 4, p. 407-15, 2011.
68. GUYATT, G. H. *et al.* GRADE guidelines: 7. Rating the quality of evidence--inconsistency. **Journal of clinical epidemiology**, v. 64, n. 12, p. 1294-1302, 2011.
69. GUYATT, G. H. *et al.* GRADE guidelines: 8. Rating the quality of evidence--indirectness. **Journal of clinical epidemiology**, v. 64, n. 12, p. 1303-1310, 2011.
70. GUYATT, G. H. *et al.* GRADE guidelines 6. Rating the quality of evidence--imprecision. **Journal of clinical epidemiology**, v. 64, n. 12, p. 1283-1293, 2011.
71. GUYATT, G. H. *et al.* GRADE guidelines: 5. Rating the quality of evidence--publication bias. **Journal of clinical epidemiology**, v. 64, n. 12, p. 1277-1282, 2011.
72. BRASIL. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. **Diretrizes metodológicas: Sistema GRADE – Manual de graduação da qualidade da evidência e força de recomendação para tomada de decisão em saúde**. Brasília, DF: Ministério da Saúde, 2014.
73. SCHÜNEMANN, H. J. *et al.* Chapter 15: Interpreting results and drawing conclusions. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.

-
74. COHEN, J. **Statistical Power Analysis in the Behavioral Sciences**. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates, Inc., 1988.
75. PAGE, M. J. *et al.* Chapter III: Reporting the review. *In*: HIGGINS, J. P. T. *et al.* (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated August 2019)**. [S. l.]: Cochrane, 2019. Disponível em: www.training.cochrane.org/handbook. Acesso em: 12 maio 2020.

Anexos

Anexo A - Tradução oficial do checklist PRISMA (www.prisma-statement.org)

Seção/tópico	N.	Item do checklist	Relatado na página n.
TÍTULO			
Título	1	Identifique o artigo como uma revisão sistemática, meta-análise ou ambas.	
RESUMO			
Resumo estruturado	2	Apresente um resumo estruturado incluindo, se aplicável: referencial teórico; objetivos; fonte de dados; critérios de elegibilidade; participantes e intervenções; avaliação do estudo e síntese dos métodos; resultados; limitações; conclusões e implicações dos achados principais; número de registro da revisão sistemática.	
INTRODUÇÃO			
Racional	3	Descreva a justificativa da revisão no contexto do que já é conhecido.	
Objetivos	4	Apresente uma afirmação explícita sobre as questões abordadas com referência a participantes, intervenções, comparações, resultados e desenho de estudos (PICOS).	
MÉTODOS			
Protocolo e registro	5	Indique se existe um protocolo de revisão, se e onde pode ser acessado (ex.: endereço eletrônico), e, se disponíveis, forneça informações sobre o registro da revisão, incluindo o número de registro.	
Critérios de elegibilidade	6	Especifique características do estudo (ex.: PICOS, extensão do seguimento) e características dos relatos (ex.: anos considerados, idioma, se é publicado) usadas como critérios de elegibilidade, apresentando justificativa.	
Fontes de informação	7	Descreva todas as fontes de informação na busca (ex.: base de dados com datas de cobertura, contato com autores para identificação de estudos adicionais) e data da última busca.	
Busca	8	Apresente a estratégia completa de busca eletrônica para pelo menos uma base de dados, incluindo os limites utilizados, de forma que possa ser repetida.	
Seleção dos estudos	9	Apresente o processo de seleção dos estudos (isto é, busca, elegibilidade, os incluídos na revisão sistemática, e, se aplicável, os incluídos na meta-análise).	
Processo de coleta de dados	10	Descreva o método de extração de dados dos artigos (ex.: formas para piloto, independente, em duplicata) e todos os processos para obtenção e confirmação de dados dos pesquisadores.	
Lista de dados	11	Liste e defina todas as variáveis obtidas dos dados (ex.: PICOS, fontes de financiamento) e quaisquer suposições ou simplificações realizadas.	

Seção/tópico	N.	Item do checklist	Relatado na página n.
Risco de viés em cada estudo	12	Descreva os métodos usados para avaliar o risco de viés em cada estudo (incluindo a especificação, se foi feita durante o estudo ou no nível de resultados), e como essa informação foi usada na análise de dados.	
Medidas de sumarização	13	Defina as principais medidas de sumarização dos resultados (ex.: risco relativo, diferença média)	
Síntese dos resultados	14	Descreva os métodos de análise dos dados e combinação de resultados dos estudos, se realizados, incluindo medidas de consistência (por exemplo, I ²) para cada meta-análise.	
Risco de viés entre estudos	15	Especifique qualquer avaliação do risco de viés que possa influenciar a evidência cumulativa (ex.: viés de publicação, relato seletivo nos estudos).	
Análises adicionais	16	Descreva métodos de análise adicional (ex.: análise de sensibilidade ou análise de subgrupos, metarregressão), se realizados, indicando quais foram pré-especificados.	
RESULTADOS			
Seleção de estudos	17	Apresente números dos estudos rastreados, avaliados para elegibilidade e incluídos na revisão, razões para exclusão em cada estágio, preferencialmente por meio de gráfico de fluxo.	
Características de estudos	18	Para cada estudo, apresente características para extração dos dados (ex.: tamanho do estudo, PICOS, período de acompanhamento) e apresente as citações.	
Risco de viés em cada estudo	19	Apresente dados sobre o risco de viés em cada estudo e, se disponível, alguma avaliação em resultados (ver item 12).	
Resultados de estudos individuais	20	Para todos os desfechos considerados (benefícios ou riscos), apresente, para cada estudo: (a) sumário simples de dados para cada grupo de intervenção; e (b) efeitos estimados e intervalos de confiança, preferencialmente por meio de gráficos de floresta.	
Síntese dos resultados	21	Apresente resultados para cada meta-análise feita, incluindo intervalos de confiança e medidas de consistência.	
Risco de viés entre os estudos	22	Apresente resultados da avaliação de risco de viés entre os estudos (ver item 15).	
Análises adicionais	23	Apresente resultados de análises adicionais, se realizadas (ex.: análise de sensibilidade ou subgrupos, metarregressão [ver item 16]).	
DISCUSSÃO			
Sumário de evidência	24	Sumarize os resultados principais, incluindo a força de evidência para cada resultado; considere sua relevância para grupos-chave (ex.: profissionais da saúde, usuários e formuladores de políticas).	
Limitações	25	Discuta limitações no nível dos estudos e dos desfechos (ex.: risco de viés) e no nível da revisão (ex.: obtenção incompleta de pesquisas identificadas, relato de viés).	
Conclusões	26	Apresente a interpretação geral dos resultados no contexto de outras evidências e implicações para futuras pesquisas.	
FINANCIAMENTO			
Financiamento	27	Descreva fontes de financiamento para revisão sistemática e outros suportes (ex.: suprimento de dados); papel dos financiadores na revisão sistemática.	

Anexo B - Principais bases de dados e seus respectivos endereços eletrônicos

Bases Principais	Endereço eletrônico
MEDLINE via PubMed	https://pubmed.ncbi.nlm.nih.gov/
Embase	https://www.embase.com/#search
Lilacs	https://lilacs.bvsalud.org/
Cochrane Library	https://www.cochranelibrary.com/
Literatura Cinzenta	
OpenGrey	http://www.opengrey.eu/
Who Library Database	http://kohahq.searo.who.int/
ProQuest	https://www.proquest.com/libraries/academic/databases/
Catálogo da Capes	http://catalogodeteses.capes.gov.br/catalogo-teses/#!/
Banco de teses da USP	https://teses.usp.br/
Bases Especializadas	
CINAHL Índice cumulativo em enfermagem e ciências afins	https://www.ebscohost.com/nursing/products/cinahl-databases/cinahl-complete
PsycINFO Behavioral & Social Sciences: Psychology	https://www.apa.org/pubs/databases/psycinfo/
PEDro Physiotherapy Evidence Database	https://www.pedro.org.au/
Adolec Base de dados de adolescentes e jovens	http://www.adolec.uerj.br/
Bdenf Base de dados de enfermagem	https://bvsenfermeria.bvsalud.org/
BVS Psicologia Base de dados em psicologia	http://www.bvs-psi.org.br/php/index
Bases Complementares	
CRD Centre for Reviews Dissemination	https://www.crd.york.ac.uk/CRDWeb/
FDA U.S. Food and Drug Administration	https://www.fda.gov/
Portal do Ministério da Saúde	http://www.saude.gov.br/index.php
Bases de Registros de Ensaios Clínicos	
clinicaltrials.gov	https://clinicaltrials.gov/
WHO International Clinical Trials Registry Platform - ICTRP	https://www.who.int/ictrp/en/
The European Union Clinical Trials register	https://www.clinicaltrialsregister.eu/
Registro Brasileiro de Ensaios Clínicos - ReBEC	http://www.ensaiosclinicos.gov.br/

MEDLINE via PubMed

Endereço eletrônico: <https://pubmed.ncbi.nlm.nih.gov/>

1. Clicar em **MeSH Database**
2. Inserir o termo referente à população (sem aspas, apóstrofes ou parênteses). Exemplo: *celiac disease*.
3. **Copiar os Entry terms** em outro documento (ex.: Microsoft Word).
4. Do lado superior direito da página, clicar em **Add to search builder**, depois, em **Search PubMed**.
5. Clicar em **Advanced**.
6. Verificar o **History and search details** – identificar o termo MeSH buscado, selecionar as **três bolinhas** da lateral esquerda e clicar em **Add query**.
7. Do lado esquerdo, selecionar a opção **Text word** e, do lado direito, copiar cada **Entry term** que está no documento do Microsoft Word correspondente à população, um de cada vez. **Copiar o primeiro Entry term** (não é necessário utilizar aspas, apóstrofo ou parênteses) no espaço **Enter a search term** e seleciona **OR** no quadrado azul da lateral direita.
8. Além dos **Entry Terms** do PubMed, podem ser incluídos outros termos livres encontrados em estudos de interesse ou em outras bases de dados que sejam referentes à população. Se houver, incluir no espaço **Enter a Search Term** e seleciona **OR** no quadrado azul da lateral direita.
9. Após finalizar os termos da população, clicar em **Search** (correspondente ao segundo quadrado azul).
10. **Realizar os mesmos passos de 1 a 9 para a intervenção**. Voltar a página inicial do PubMed – MeSH Database. Não é sempre que haverá um MeSH para a intervenção, ela pode aparecer como *supplementary concept* (estará escrito “*supplementary concept*” na frente do termo buscado). Se não houver nenhum dos dois, os autores precisam buscar termos livres nos estudos que identifiquem a intervenção, por exemplo: nome comercial da intervenção, entre outros.
11. Após buscar os termos da intervenção, os autores devem voltar ao **Advanced**, identificar no histórico a busca da população, selecionar nas três bolinhas da lateral esquerda a opção **Add query** e, em seguida, identificar a busca da intervenção no histórico e selecionar nas três bolinhas da lateral esquerda a opção **Add with AND**. Clicar em **Search**.
12. A busca estará finalizada. Copie a busca completa no espaço que estará do lado esquerdo do quadrado azul com a palavra **Search**. No Microsoft Word, anote o número de resultados encontrados e a data de busca, preferencialmente em um quadro.
13. Para extrair os resultados, clique em **Send to**, selecione **Citation manager**, depois selecione a opção **All results**. Clique em **Create file**. Salve o arquivo baixado.

(continua)

(continuação)

Embase

Endereço eletrônico: <https://www.embase.com/#search>

1. Clicar em **Emtree**.
2. Escrever o termo referente à população. Exemplo: *celiac disease*.
3. **Copiar os Synonyms** do termo em um arquivo do Microsoft Word.
4. Selecionar **Explode**, clicar em **Add to query builder**. Clicar em **Search**.
5. Voltar no **Search - Quick** e colocar termo por termo dos **Emtree**. Os termos devem ser colocados entre apóstrofes (ex.: '*adult celiac disease*'). Selecionar o **All fields** do lado esquerdo e o **OR**.

Outra maneira é fazer pelo **Results**. Colocar termo por termo, depois clicar em **Combine** com **OR** todos os termos referente à população.

Quando tiver todos os termos da população, ir ao lado do **Search**, **selecionar todos os ícones do Mapping** e no ícone **Sources**, selecionar **Embase** e realizar a busca, pois os resultados do MEDLINE também estão disponíveis; ou seja, serão gerados muitos estudos duplicados.

6. Seguir os mesmos passos para a intervenção (1 ao 5).
7. Combinar os termos da população com os termos da intervenção, clicando em **Combine** e selecionando **AND**.
8. Copiar a estratégia de busca completa no Microsoft Word e a data de busca.
9. Clicar no número correspondente à busca completa, ir à caixa **Selected**, selecionar o valor correspondente a **todos os resultados**, clicar em **Export**, formato **RIS format (Mendeley, Endnote)**. Salvar o documento exportado.

(continua)

(continuação)

Lilacs via Portal Regional da BVS

1. Buscar os termos da população e da intervenção no Descritores em Ciências da Saúde (DeCS). Buscar em todos os idiomas disponíveis (português, inglês, espanhol). Copiar os descritores de assunto e os sinônimos em um arquivo do Microsoft Word.

Endereço eletrônico do DeCS: http://decs.bvs.br/cgi-bin/wxis1660.exe/decserver/?IscScript=../cgi-bin/decserver/decserver.xis&interface_language=p&previous_page=homepage&previous_task=NULL&task=start

2. **Entrar no endereço eletrônico do Portal Regional da BVS:** <https://pesquisa.bvsalud.org/portal/advanced/?lang=pt>
3. Na **busca avançada**, colocar todos os termos encontrados no DeCS em relação à população. Os descritores dos três idiomas devem ser identificados como descritor de assunto (no seletor do lado direito), e os sinônimos devem ser identificados como título, resumo, assunto. Combinar os termos com **OR**. Podem ser adicionados também os códigos da “categoria” do DeCS, inserindo o termo “ex” no início do código e o “\$” no final (ex.: exG07.203.650.240.265\$). Estes devem ser combinados com OR com os outros termos referentes à população.
4. **Pesquisar.**
5. Copiar a estratégia de busca fornecida no pesquisador em um documento do Microsoft Word.
6. Voltar na **busca avançada**, repetir os passos 3 e 4 em relação à intervenção e combinar com **OR**.
7. Depois repetir o item 5 para intervenção.
8. Unir, no Word, a estratégia de busca da população e da intervenção, adicionando um parêntese no início e no final de cada estratégia, e unir com **AND**. Colocar a nova estratégia no pesquisador ao lado da lupa azul.
9. Após disponibilizados os resultados, **selecionar a base de dados Lilacs** do lado esquerdo.
10. Exportar todos os resultados.

(continua)

(continuação)

Cochrane Library

Endereço eletrônico: <https://www.cochranelibrary.com/>.

1. Clicar em **Advanced search**.
2. Clicar em **Medical terms (MeSH)**.
3. Inserir o termo referente à população, clicar em **Look up**.
4. Copiar, em um documento do Microsoft Word, os termos que se encontram no **Exact term match**.
5. Selecionar **Explode all trees**, depois clicar em **Add to search manager**.
6. Adicionar no **Search manager** os sinônimos da população (um por um, sem aspas ou apóstrofos), depois colocar #1 OR #2 OR #3, e assim por diante, para todos os termos da população.
7. Repetir os passos 2 a 6 para intervenção.
8. Unir os termos da população com os da intervenção com **AND**.
9. Clicar no número dos resultados. Clicar em **Select all**, depois clicar em **Export selected citations**. Salvar, no formato RIS (**Reference Manager**), o documento baixado.

Anexo D - Exemplos de estratégias de busca

Exemplos de estratégias de busca na base de dados MEDLINE via PubMed

MEDLINE via PubMed

A. Busca sensível com População e Intervenção

Data da busca: 15/7/2020

Resultado: 5.181 referências

População: Celiac Disease

((((((((((((((((((("Celiac Disease"[Mesh]) OR (Disease, Celiac)) OR (Gluten Enteropathy)) OR (Enteropathies, Gluten)) OR (Enteropathy, Gluten)) OR (Gluten Enteropathies)) OR (Gluten-Sensitive Enteropathy)) OR (Enteropathies, Gluten-Sensitive)) OR (Enteropathy, Gluten-Sensitive)) OR (Gluten Sensitive Enteropathy)) OR (Gluten-Sensitive Enteropathies)) OR (Sprue, Celiac)) OR (Sprue, Nontropical)) OR (Nontropical Sprue)) OR (Celiac Sprue)) OR (Sprue)) OR (Coeliac disease)))

AND

Intervenção: Diet, Gluten-Free

((((((("Diet, Gluten-Free"[Mesh]) OR (Diet, Gluten Free)) OR (Gluten-Free Diet)) OR (Diets, Gluten-Free)) OR (Gluten Free Diet)) OR (Gluten-Free Diets))

B. Busca com População, Intervenção e Filtro Validado para ECR (sensível)

Data da busca: 15/7/2020

Resultado: 1.059 referências

População: Celiac Disease

((((((((((((((((((("Celiac Disease"[Mesh]) OR (Disease, Celiac)) OR (Gluten Enteropathy)) OR (Enteropathies, Gluten)) OR (Enteropathy, Gluten)) OR (Gluten Enteropathies)) OR (Gluten-Sensitive Enteropathy)) OR (Enteropathies, Gluten-Sensitive)) OR (Enteropathy, Gluten-Sensitive)) OR (Gluten Sensitive Enteropathy)) OR (Gluten-Sensitive Enteropathies)) OR (Sprue, Celiac)) OR (Sprue, Nontropical)) OR (Nontropical Sprue)) OR (Celiac Sprue)) OR (Sprue)) OR (Coeliac disease)))

AND

Intervenção: Diet, Gluten-Free

((((((("Diet, Gluten-Free"[Mesh]) OR (Diet, Gluten Free)) OR (Gluten-Free Diet)) OR (Diets, Gluten-Free)) OR (Gluten Free Diet)) OR (Gluten-Free Diets))

AND

Delineamento de estudo: Ensaio Clínico Randomizado

((randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR drug therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab] NOT (animals [mh] NOT humans [mh])))

C. Busca com População, Intervenção e Filtro Validado para ECR (sensível e específico)

Data da busca: 15/7/2020

Resultado: 260 referências

População: Celiac Disease

((((((((((((((("Celiac Disease"[Mesh]) OR (Disease, Celiac)) OR (Gluten Enteropathy)) OR (Enteropathies, Gluten)) OR (Enteropathy, Gluten)) OR (Gluten Enteropathies)) OR (Gluten-Sensitive Enteropathy)) OR (Enteropathies, Gluten-Sensitive)) OR (Enteropathy, Gluten-Sensitive)) OR (Gluten Sensitive Enteropathy)) OR (Gluten-Sensitive Enteropathies)) OR (Sprue, Celiac)) OR (Sprue, Nontropical)) OR (Nontropical Sprue)) OR (Celiac Sprue)) OR (Sprue)) OR (Coeliac disease)))

AND

Intervenção: Diet, Gluten-Free

(((((("Diet, Gluten-Free"[Mesh]) OR (Diet, Gluten Free)) OR (Gluten-Free Diet)) OR (Diets, Gluten-Free)) OR (Gluten Free Diet)) OR (Gluten-Free Diets))

AND

Delineamento de estudo: Ensaio Clínico Randomizado

((randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR placebo[tiab] OR clinical trials as topic[mesh:noexp] OR randomly[tiab] OR trial[ti] NOT (animals[mh] NOT humans [mh])))

Exemplos de estratégias de busca na base de dados Embase

Embase

A. Busca sensível com População e Intervenção

Data da busca: 15/7/2020

Resultado: 8.417 referências

População: Celiac Disease

('celiac disease'/exp OR 'celiac disease' OR 'adult celiac disease'/exp OR 'adult celiac disease' OR 'adult coeliac disease'/exp OR 'adult coeliac disease' OR 'celiac sprue'/exp OR 'celiac sprue' OR 'celiac syndrome'/exp OR 'celiac syndrome' OR 'coeliac disease'/exp OR 'coeliac disease' OR 'coeliac sprue'/exp OR 'coeliac sprue' OR 'coeliac syndrome'/exp OR 'coeliac syndrome' OR 'coeliaky'/exp OR 'coeliaky' OR 'disease, celiac'/exp OR 'disease, celiac' OR 'disease, coeliac'/exp OR 'disease, coeliac' OR 'gee herter disease'/exp OR 'gee herter disease' OR 'gee thaysen disease'/exp OR 'gee thaysen disease' OR 'gluten enteropathy'/exp OR 'gluten enteropathy' OR 'gluten induced enteropathy'/exp OR 'gluten induced enteropathy' OR 'gluten intolerance'/exp OR 'gluten intolerance' OR 'gluten sensitive enteropathy'/exp OR 'gluten sensitive enteropathy' OR 'huebner herter disease'/exp OR 'huebner herter disease' OR 'infantile celiac disease'/exp OR 'infantile celiac disease' OR 'infantile coeliac disease'/exp OR 'infantile coeliac disease' OR 'infantile indigenus sprue'/exp OR 'infantile indigenus sprue' OR 'infantilism, intestinal'/exp OR 'infantilism, intestinal' OR 'intestinal infantilism'/exp OR 'intestinal infantilism' OR 'nontropical sprue'/exp OR 'nontropical sprue' OR 'sprue, indigenus'/exp OR 'sprue, indigenus' OR 'sprue, nontropical'/exp OR 'sprue, nontropical') AND [embase]/lim

AND

Intervenção: Gluten Free Diet

('gluten free diet'/exp OR 'gluten free diet' OR 'diet, gluten-free'/exp OR 'diet, gluten-free' OR 'diet, gluten free'/exp OR 'diet, gluten free') AND [embase]/lim

(continua)

(continuação)

B. Busca com População, Intervenção e Filtro Validado para ECR

Data da busca: 15/7/2020

Resultado: 637 referências

População: Celiac Disease

('celiac disease'/exp OR 'celiac disease' OR 'adult celiac disease'/exp OR 'adult celiac disease' OR 'adult coeliac disease'/exp OR 'adult coeliac disease' OR 'celiac sprue'/exp OR 'celiac sprue' OR 'celiac syndrome'/exp OR 'celiac syndrome' OR 'coeliac disease'/exp OR 'coeliac disease' OR 'coeliac sprue'/exp OR 'coeliac sprue' OR 'coeliac syndrome'/exp OR 'coeliac syndrome' OR 'coeliaky'/exp OR 'coeliaky' OR 'disease, celiac'/exp OR 'disease, celiac' OR 'disease, coeliac'/exp OR 'disease, coeliac' OR 'gee herter disease'/exp OR 'gee herter disease' OR 'gee thaysen disease'/exp OR 'gee thaysen disease' OR 'gluten enteropathy'/exp OR 'gluten enteropathy' OR 'gluten induced enteropathy'/exp OR 'gluten induced enteropathy' OR 'gluten intolerance'/exp OR 'gluten intolerance' OR 'gluten sensitive enteropathy'/exp OR 'gluten sensitive enteropathy' OR 'huebner herter disease'/exp OR 'huebner herter disease' OR 'infantile celiac disease'/exp OR 'infantile celiac disease' OR 'infantile coeliac disease'/exp OR 'infantile coeliac disease' OR 'infantile indigenus sprue'/exp OR 'infantile indigenus sprue' OR 'infantilism, intestinal'/exp OR 'infantilism, intestinal' OR 'intestinal infantilism'/exp OR 'intestinal infantilism' OR 'nontropical sprue'/exp OR 'nontropical sprue' OR 'sprue, indigenus'/exp OR 'sprue, indigenus' OR 'sprue, nontropical'/exp OR 'sprue, nontropical') AND [embase]/lim

AND

Intervenção: Gluten Free Diet

('gluten free diet'/exp OR 'gluten free diet' OR 'diet, gluten-free'/exp OR 'diet, gluten-free' OR 'diet, gluten free'/exp OR 'diet, gluten free') AND [embase]/lim

AND

Delineamento de estudo: Ensaio Clínico Randomizado

('crossover procedure':de OR 'double-blind procedure':de OR 'randomized controlled trial':de OR 'single-blind procedure':de OR random*:de,ab,ti OR factorial*:de,ab,ti OR crossover*:de,ab,ti OR ((cross NEXT/1 over*):de,ab,ti) OR placebo*:de,ab,ti OR ((doubl* NEAR/1 blind*):de,ab,ti) OR ((singl* NEAR/1 blind*):de,ab,ti) OR assign*:de,ab,ti OR allocat*:de,ab,ti OR volunteer*:de,ab,ti) AND [embase]/lim

(continua)

(continuação)

C. Busca com População, Intervenção e Filtro Validado para ECR (sensível)

Data da busca: 15/7/2020

Resultado: 2.152 referências

População: Celiac Disease

('celiac disease'/exp OR 'celiac disease' OR 'adult celiac disease'/exp OR 'adult celiac disease' OR 'adult coeliac disease'/exp OR 'adult coeliac disease' OR 'celiac sprue'/exp OR 'celiac sprue' OR 'celiac syndrome'/exp OR 'celiac syndrome' OR 'coeliac disease'/exp OR 'coeliac disease' OR 'coeliac sprue'/exp OR 'coeliac sprue' OR 'coeliac syndrome'/exp OR 'coeliac syndrome' OR 'coeliaky'/exp OR 'coeliaky' OR 'disease, celiac'/exp OR 'disease, celiac' OR 'disease, coeliac'/exp OR 'disease, coeliac' OR 'gee herter disease'/exp OR 'gee herter disease' OR 'gee thaysen disease'/exp OR 'gee thaysen disease' OR 'gluten enteropathy'/exp OR 'gluten enteropathy' OR 'gluten induced enteropathy'/exp OR 'gluten induced enteropathy' OR 'gluten intolerance'/exp OR 'gluten intolerance' OR 'gluten sensitive enteropathy'/exp OR 'gluten sensitive enteropathy' OR 'huebner herter disease'/exp OR 'huebner herter disease' OR 'infantile celiac disease'/exp OR 'infantile celiac disease' OR 'infantile coeliac disease'/exp OR 'infantile coeliac disease' OR 'infantile indigenus sprue'/exp OR 'infantile indigenus sprue' OR 'infantilism, intestinal'/exp OR 'infantilism, intestinal' OR 'intestinal infantilism'/exp OR 'intestinal infantilism' OR 'nontropical sprue'/exp OR 'nontropical sprue' OR 'sprue, indigenus'/exp OR 'sprue, indigenus' OR 'sprue, nontropical'/exp OR 'sprue, nontropical') AND [embase]/lim

AND

Intervenção: Gluten Free Diet

('gluten free diet'/exp OR 'gluten free diet' OR 'diet, gluten-free'/exp OR 'diet, gluten-free' OR 'diet, gluten free'/exp OR 'diet, gluten free') AND [embase]/lim

AND

Delineamento de estudo: Ensaio Clínico Randomizado

(random*:ab,ti OR ((clinical NEXT/1 trial*):de,ab,ti) OR 'health care quality'/exp OR 'health care quality') AND [embase]/lim

D. Busca com População, Intervenção e Filtro Validado para ECR (específico)

Data da busca: 15/7/2020

Resultado: 261 referências

População: Celiac Disease

('celiac disease'/exp OR 'celiac disease' OR 'adult celiac disease'/exp OR 'adult celiac disease' OR 'adult coeliac disease'/exp OR 'adult coeliac disease' OR 'celiac sprue'/exp OR 'celiac sprue' OR 'celiac syndrome'/exp OR 'celiac syndrome' OR 'coeliac disease'/exp OR 'coeliac disease' OR 'coeliac sprue'/exp OR 'coeliac sprue' OR 'coeliac syndrome'/exp OR 'coeliac syndrome' OR 'coeliaky'/exp OR 'coeliaky' OR 'disease, celiac'/exp OR 'disease, celiac' OR 'disease, coeliac'/exp OR 'disease, coeliac' OR 'gee herter disease'/exp OR 'gee herter disease' OR 'gee thaysen disease'/exp OR 'gee thaysen disease' OR 'gluten enteropathy'/exp OR 'gluten enteropathy' OR 'gluten induced enteropathy'/exp OR 'gluten induced enteropathy' OR 'gluten intolerance'/exp OR 'gluten intolerance' OR 'gluten sensitive enteropathy'/exp OR 'gluten sensitive enteropathy' OR 'huebner herter disease'/exp OR 'huebner herter disease' OR 'infantile celiac disease'/exp OR 'infantile celiac disease' OR 'infantile coeliac disease'/exp OR 'infantile coeliac disease' OR 'infantile indigenus sprue'/exp OR 'infantile indigenus sprue' OR 'infantilism, intestinal'/exp OR 'infantilism, intestinal' OR 'intestinal infantilism'/exp OR 'intestinal infantilism' OR 'nontropical sprue'/exp OR 'nontropical sprue' OR 'sprue, indigenus'/exp OR 'sprue, indigenus' OR 'sprue, nontropical'/exp OR 'sprue, nontropical') AND [embase]/lim

AND

Intervenção: Gluten Free Diet

('gluten free diet'/exp OR 'gluten free diet' OR 'diet, gluten-free'/exp OR 'diet, gluten-free' OR 'diet, gluten free'/exp OR 'diet, gluten free') AND [embase]/lim

AND

Delineamento de estudo: Ensaio Clínico Randomizado

((double NEXT/1 blind*):de,ab,ti) OR placebo*:ab,ti OR blind*:ab,ti) AND [embase]/lim

(continua)

(continuação)

E. Busca com População, Intervenção e Filtro Validado para ECR (sensível e específico)

Data da busca: 15/7/2020

Resultado: 443 referências

População: Celiac Disease

('celiac disease'/exp OR 'celiac disease' OR 'adult celiac disease'/exp OR 'adult celiac disease' OR 'adult coeliac disease'/exp OR 'adult coeliac disease' OR 'celiac sprue'/exp OR 'celiac sprue' OR 'celiac syndrome'/exp OR 'celiac syndrome' OR 'coeliac disease'/exp OR 'coeliac disease' OR 'coeliac sprue'/exp OR 'coeliac sprue' OR 'coeliac syndrome'/exp OR 'coeliac syndrome' OR 'coeliaky'/exp OR 'coeliaky' OR 'disease, celiac'/exp OR 'disease, celiac' OR 'disease, coeliac'/exp OR 'disease, coeliac' OR 'gee herter disease'/exp OR 'gee herter disease' OR 'gee thaysen disease'/exp OR 'gee thaysen disease' OR 'gluten enteropathy'/exp OR 'gluten enteropathy' OR 'gluten induced enteropathy'/exp OR 'gluten induced enteropathy' OR 'gluten intolerance'/exp OR 'gluten intolerance' OR 'gluten sensitive enteropathy'/exp OR 'gluten sensitive enteropathy' OR 'huebner herter disease'/exp OR 'huebner herter disease' OR 'infantile celiac disease'/exp OR 'infantile celiac disease' OR 'infantile coeliac disease'/exp OR 'infantile coeliac disease' OR 'infantile indigenous sprue'/exp OR 'infantile indigenous sprue' OR 'infantilism, intestinal'/exp OR 'infantilism, intestinal' OR 'intestinal infantilism'/exp OR 'intestinal infantilism' OR 'nontropical sprue'/exp OR 'nontropical sprue' OR 'sprue, indigenous'/exp OR 'sprue, indigenous' OR 'sprue, nontropical'/exp OR 'sprue, nontropical') AND [embase]/lim

AND

Intervenção: Gluten Free Diet

('gluten free diet'/exp OR 'gluten free diet' OR 'diet, gluten-free'/exp OR 'diet, gluten-free' OR 'diet, gluten free'/exp OR 'diet, gluten free') AND [embase]/lim

AND

Delineamento de estudo: Ensaio Clínico Randomizado

(random*:ab,ti OR placebo*:de,ab,ti OR ((double NEXT/1 blind*):ab,ti)) AND [embase]/lim

Exemplos de estratégias de busca na base de dados Cochrane Library

Cochrane Library

A. Busca sensível com População e Intervenção

Data da busca: 15/7/2020

Resultado: 372 referências

#1	MeSH descriptor: [Celiac Disease] explode all trees	339
#2	Gluten-Sensitive Enteropathies	0
#3	Gluten Sensitive Enteropathy	3
#4	Enteropathy, Gluten-Sensitive	2
#5	Enteropathies, Gluten-Sensitive	0
#6	Gluten-Sensitive Enteropathy	2
#7	Gluten Enteropathies	3
#8	Celiac Sprue	15
#9	Sprue, Celiac	15
#10	Gluten Enteropathy	21
#11	Sprue, Nontropical	1
#12	Nontropical Sprue	1
#13	Enteropathy, Gluten	21
#14	Enteropathies, Gluten	3
#15	Disease, Celiac	961
#16	Sprue	30
#17	(OR #1-#16)0	
#18	MeSH descriptor: [Diet, Gluten-Free] explode all trees	81
#19	Gluten-Free Diets	63
#20	Diet, Gluten Free	509
#21	Gluten Free Diet	509
#22	Gluten-Free Diet	487
#23	Diets, Gluten-Free	63
#24	(OR #18-#23)	513
#25	#17 AND #24	372

Exemplos de estratégias de busca na base de dados Lilacs via Portal Regional da BVS

Lilacs via Portal Regional da BVS

A. Busca sensível com População e Intervenção

Data da busca: 15/7/2020

Resultado: 137 referências

População: Celiac Disease

(mh:(Doença celíaca)) OR (mh:(Enfermedad Celíaca)) OR (mh:(Celiac Disease)) OR (tw:(Enteropatia Glúten Induzida)) OR (tw:(Espru Celíaco)) OR (tw:(Espru não Tropical)) OR (tw:(Estearreia)) OR (tw:(Enteropatía por Gluten)) OR (tw:(Esprue Celíaco)) OR (tw:(Esprue no Tropical)) OR (tw:(Celiac Sprue)) OR (tw:(Disease, Celiac)) OR (tw:(Enteropathies, Gluten)) OR (tw:(Enteropathies, Gluten-Sensitive)) OR (tw:(Enteropathy, Gluten-Sensitive)) OR (tw:(Gluten Enteropathies)) OR (tw:(Gluten Enteropathy)) OR (tw:(Gluten Sensitive Enteropathy)) OR (tw:(Gluten-Sensitive Enteropathies)) OR (tw:(Gluten-Sensitive Enteropathy)) OR (tw:(Nontropical Sprue)) OR (tw:(Sprue)) OR (tw:(Sprue, Celiac)) OR (tw:(Sprue, Nontropical)) OR (tw:(exC06.405.469.637.250\$)) OR (tw:(exC18.452.603.250\$))

AND

Intervenção: Diet, Gluten-Free

(mh:(diet, gluten-free)) OR (mh:(dieta sin gluten)) OR (mh:(dieta livre de glúten)) OR (tw:(Diet, Gluten Free)) OR (tw:(Diets, Gluten-Free)) OR (tw:(Gluten Free Diet)) OR (tw:(Gluten-Free Diet)) OR (tw:(Gluten-Free Diets)) OR (tw:(exE02.642.249.265\$)) OR (tw:(exG07.203.650.240.265\$))

Anexo E – Tradução não validada da ferramenta RoB-2 para avaliação do risco de viés de ensaios clínicos randomizados

Domínio 1: risco de viés no processo de randomização

Questões	Comentários	Opções de resposta
1.1 A sequência de alocação dos participantes foi aleatória?		Sim (S) / Parcialmente Sim (PS) / Parcialmente Não (PN) / Não (N) / Nenhuma Informação (NI)
1.2 Foi mantido o sigilo de alocação dos participantes até eles serem recrutados e alocados para as intervenções?		S / PS / PN / N / NI
1.3 As diferenças na linha de base entre os grupos de intervenção sugeriram um problema com o processo de randomização?		S / PS / PN / N / NI
<i>Julgamento do risco de viés</i>		Baixo risco / Alto risco / Algumas preocupações

Domínio 2: risco de viés devido a desvios das intervenções pretendidas

Questões	Comentários	Opções de resposta
2.1 Os participantes tinham conhecimento em relação às intervenções fornecidas a eles durante o estudo?		S / PS / PN / N / NI
2.2 Os cuidadores ou as pessoas que estavam entregando as intervenções tinham conhecimento em relação às intervenções fornecidas aos participantes durante o estudo?		S / PS / PN / N / NI
2.3 Se as perguntas 2.1 e 2.2 foram respondidas como S / PS / NI: os desvios da intervenção pretendida surgiram devido ao contexto do estudo?		NA / S / PS / PN / N / NI
2.4 Se a pergunta 2.3 foi respondida como S / PS: os desvios provavelmente afetaram o desfecho?		NA / S / PS / PN / N / NI
2.5 Se a pergunta 2.4 foi respondida como S / PS: os desvios da intervenção pretendida foram balanceados entre os grupos?		NA / S / PS / PN / N / NI
2.6 Foi utilizada uma análise apropriada para estimar o efeito da intervenção alocada?		S / PS / PN / N / NI
2.7 Se a pergunta 2.6 foi respondida como N / PN / NI: existiu potencial para um impacto substancial (no resultado) da falha de analisar os participantes no grupo em que foram randomizados?		NA / S / PS / PN / N / NI
<i>Julgamento do risco de viés</i>		Baixo risco / Alto risco / Algumas preocupações

Domínio 3: risco de viés devido a dados perdidos do desfecho

Questões	Comentários	Opções de resposta
3.1 Os dados para esse desfecho estão disponíveis para todos ou quase todos os participantes randomizados?		S / PS / PN / N / NI
3.2 Se a pergunta 3.1 foi respondida como N / PN / NI: existe evidência de que o resultado não foi enviesado pelos dados perdidos do desfecho?		NA / S / PS / PN / N
3.3 Se a pergunta 3.2 foi respondida como N / PN: a perda de dados do desfecho poderia depender de seu valor verdadeiro?		NA / S / PS / PN / N / NI
3.4 Se a pergunta 3.3 foi respondida como S / PS / NI: é provável que a perda de dados do desfecho dependa do seu valor verdadeiro?		NA / S / PS / PN / N / NI
<i>Julgamento do risco de viés</i>		Baixo risco / Alto risco / Algumas preocupações

Domínio 4: risco de viés na mensuração do desfecho

Questões	Comentários	Opções de resposta
4.1 O método de mensuração do desfecho foi inapropriado?		S / PS / PN / N / NI
4.2 A mensuração do desfecho pode ter diferido entre os grupos de intervenção?		S / PS / PN / N / NI
4.3 Se as perguntas 4.1 e 4.2 foram respondidas como N / PN / NI: os avaliadores do desfecho estavam cientes da intervenção recebida pelos participantes do estudo?		NA / S / PS / PN / N / NI
4.4 Se a pergunta 4.3 foi respondida como S / PS / NI: a avaliação do desfecho pode ter sido influenciada pelo conhecimento da intervenção recebida?		NA / S / PS / PN / N / NI
4.5 Se a pergunta 4.4 foi respondida como S / PS / NI: é provável que a avaliação do desfecho tenha sido influenciada pelo conhecimento da intervenção recebida?		NA / S / PS / PN / N / NI
<i>Julgamento do risco de viés</i>		Baixo risco / Alto risco / Algumas preocupações

Domínio 5: risco de viés na seleção do resultado reportado

Questões	Comentários	Opções de resposta
5.1 Os dados que produziram o resultado foram analisados de acordo com o plano de análise pré-especificado, o qual foi finalizado antes de os dados não cegos dos desfechos estarem disponíveis para análise?		S / PS / PN / N / NI
O resultado numérico avaliado provavelmente foi selecionado, baseado nos resultados de...		
5.2 ... múltiplas mensurações elegíveis (ex.: escalas, definições, pontos de seguimento) do desfecho dentro do domínio desfecho?		S / PS / PN / N / NI
5.3 ... múltiplas análises elegíveis do dado?		S / PS / PN / N / NI
<i>Julgamento do risco de viés</i>		Baixo risco / Alto risco / Algumas preocupações

Risco de viés geral

Critérios		Julgamento do risco de viés
O estudo é julgado de ser baixo risco de viés para todos os domínios para esse resultado.		Baixo risco
O estudo é julgado de ter algumas preocupações para ao menos um domínio para esse resultado, mas não é julgado de ser alto risco para nenhum domínio.		Algumas preocupações
O estudo é julgado de ser alto risco de viés para pelo menos um domínio para esse resultado. OU O estudo é julgado de ter algumas preocupações para múltiplos domínios em uma maneira que substancialmente diminui a confiança no resultado.		Alto risco Alto risco

Fonte: Ferramenta RoB-2 da Cochrane (STERNE, J. A. C. *et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. **BMJ**, v. 366, l4898, 2019.)

Conte-nos o que pensa sobre esta publicação. [Clique aqui](#) e responda a pesquisa.



PROADI SUS
Programa de Apoio ao Desenvolvimento
Institucional do Sistema Único de Saúde

DISQUE
SAÚDE
136

Biblioteca Virtual em Saúde
do Ministério da Saúde
bvsms.saude.gov.br



OSWALDO CRUZ
HOSPITAL ALEMÃO

REBRATS



MINISTÉRIO DA
SAÚDE

**Governo
Federal**