

**BASES DE DADOS E DEMAIS FONTES DE INFORMAÇÃO EM REVISÕES  
SISTEMÁTICAS: recomendações práticas**

***DATABASES AND OTHER SOURCES OF INFORMATION IN SYSTEMATIC RE-  
VIEWS: practical recommendations***

 Wánderon Cássio Oliveira Araújo<sup>1</sup>

 Fernanda Machado Lopes<sup>2</sup>

 Natália Martins Dias<sup>3</sup>

<sup>1</sup> Doutorando em Ciência da Informação na Universidade Federal de Santa Catarina (UFSC). Bibliotecário na Universidade Federal do Ceará na Biblioteca de Ciências da Saúde.

**E-mail:** wcassio@ufc.br

<sup>2</sup> Doutora em Psicologia. Professora do Programa de Pós-Graduação em Psicologia da Universidade Federal de Santa Catarina (UFSC).

**E-mail:** femlopes23@gmail.com

<sup>3</sup> Doutora em Distúrbios do Desenvolvimento. Professora do Programa de Pós-Graduação em Psicologia da Universidade Federal de Santa Catarina (UFSC).

**E-mail:** natalia.m.dias@ufsc.br

<sup>4</sup> Doutora em Psicologia. Professora do Programa de Pós-Graduação em Psicologia da Universidade Federal de Santa Catarina (UFSC).

**E-mail:** andreabs@gmail.com

<sup>5</sup> Doutoranda, mestre e graduada em Psicologia no Programa de Pós-Graduação em Psicologia da Universidade Federal de Santa Catarina (UFSC).

**E-mail:** marianaluiza\_b@hotmail.com

<sup>6</sup> Doutora em Engenharia de Produção. Professora do Programa de Pós-Graduação em Psicologia da Universidade Federal de Santa Catarina (UFSC).

 Andréa Barbará da Silva Bousfield<sup>4</sup>

 Mariana Luiza Becker da Silva<sup>5</sup>

 Andrea Valéria Steil<sup>6</sup>

**E-mail:** andrea.steil@ufsc.br



**ACESSO ABERTO**

**Copyright:** Esta obra está licenciada com uma Licença Creative Commons Atribuição 4.0 Internacional. 

**Conflito de interesses:** Os autores declaram que não há conflito de interesses.

**Financiamento:** Não há.

**Declaração de Disponibilidade dos dados:** Todos os dados relevantes estão disponíveis neste artigo.

**Recebido em:** 31 jan. 2024.

**Aceito em:** 31 jun. 2024.

**Publicado em:** 04 nov. 2024.

**Como citar este artigo:**

ARAÚJO, W. C. O.; LOPES, F. M.; DIAS, N. M.; BOUSFIELD, A. B. S.; SILVA, M. L. B.; STEIL, A. V. Bases de dados e demais fontes de informação em revisões sistemáticas: recomendações práticas. **Informação em Pauta**, Fortaleza, v. 9, p. 1-19, 2024.

**RESUMO**

Bases de dados bibliográficas são a principal fonte de informação para a identificação e a recuperação de estudos primários a serem sintetizados em revisões sistemáticas (RS). A partir da análise das fontes de informação utilizadas em 305 RS em psicologia,

o objetivo deste artigo é explicitar como as fontes de informação têm sido reportadas em RS em psicologia e apresentar soluções práticas para as lacunas e incoerências identificadas. O artigo apresenta uma classificação das fontes de informação acadêmicas bibliográficas, define base de dados bibliográfica e demonstra como nomear corretamente as fontes de informação em RS. Também discorre sobre a quantidade de bases em RS, sobre a construção de estratégias de busca, sobre o uso de vocabulário controlado e sobre a explosão em bases de dados. Por fim, o artigo discute a questão da literatura cinza, o uso de fontes complementares, a questão da regionalidade das fontes de informação e o papel do bibliotecário nas RS. Espera-se que as questões discutidas e as recomendações práticas desse artigo contribuam para o aperfeiçoamento das RS realizadas por pesquisadores em todas as áreas do conhecimento.

**Palavras-chave:** revisão sistemática; fontes de informação; bases de dados; recuperação da informação.

## ABSTRACT

Bibliographic databases are the main source of information for identifying and retrieving primary studies to be synthesized in systematic reviews (SR). After analyzing the information sources used in 305 SR in psychology, this article explains how the information sources have been reported and presents practical solutions to the gaps and inconsistencies identified. The article classifies academic bibliographic information sources, defines bibliographic databases, and demonstrates how to name information sources in SR correctly. It also discusses the number of databases in SR, the construction of search strategies, the use of controlled vocabulary, and the databases burst. Finally, the article discusses gray literature, the use of complementary sources, the regionality of information sources, and the librarian's role in SR. It is hoped that the issues discussed, and the practical recommendations of this article contribute to improving SR performed by researchers in all areas of knowledge.

**Keywords:** systematic review; information sources; database; information retrieval.

---

## 1 INTRODUÇÃO

Revisões sistemáticas (RS) começaram a ser publicadas em grande número na área da saúde a partir de 1980 (Hansen; Trifkovic, 2013). Nas últimas duas décadas, as RS ganharam ainda mais espaço, volume e importância em diferentes áreas do conhecimento. Na PubMed, uma das fontes de informação mais relevantes em Ciências da Saúde, até o ano 2000, tinha um total 2.274 RS indexadas. Vinte anos depois, no ano de 2020, esse número saltou para 163.237 RS indexadas. Somente no ano de 2022, nesta fonte de informação, foram indexadas 38.503 RS sendo este o período com maior indexação de publicações de RS da série histórica. Tais números denotam o constante crescimento deste tipo de investigação.

Uma RS utiliza procedimentos explícitos e sistemáticos para coletar, combinar e sintetizar os resultados de estudos primários que respondam a uma questão claramente formulada (Higgins et al., 2021). Configura-se em um processo lógico e linear em que cada uma de suas etapas está, obrigatoriamente, conectada às demais (Purssell; McCrae, 2020). Seu objetivo é gerar uma resposta objetiva, embasada empiricamente, para uma questão de pesquisa a partir da identificação, síntese e avaliação de todas as evidências disponíveis (Patole, 2021). Esse tipo de revisão pode ser aplicado a diferentes tipos de questões de pesquisa (Newman; Gough, 2020) e em diferentes campos do saber (Patole, 2021). Quando consideramos a pirâmide de evidências, as revisões sistemáticas estão alocadas em seu topo (Murad et al., 2016), ou seja, quando uma RS é conduzida adequadamente, as respostas encontradas são evidências científicas altamente confiáveis e que geralmente são aplicadas, de forma prática, na sua área do conhecimento.

As RS têm exercido importante papel no desenvolvimento científico de diferentes áreas do conhecimento por

sintetizar evidências de estudos já realizados, apontar lacunas de pesquisa, ajudar a organizar princípios teóricos, conectar abordagens teóricas com resultados práticos, subsidiar decisões clínicas, entre outros. O uso de RS também tem crescido em função de sua associação com o movimento da prática baseada em evidências na área da saúde, uma vez que um de seus pilares é a busca da melhor evidência de pesquisa disponível, integrada à experiência clínica e adaptada às características do cliente/paciente (Melnik; Souza; Carvalho, 2014).

No entanto, a profusão quantitativa de RS não é proporcional à qualidade metodológica esperada (Park et al., 2022; Ramasamy, 2022; Steil et al., 2022). Pesquisadores da área da saúde, educação, engenharia, computação e psicologia já indicaram que a qualidade de uma revisão sistemática está associada à aderência ao método de sua realização (Borrego; Foster; Froyd, 2015; Grainger et al., 2020; Higgins; Green, 2011; Siddaway et al., 2019). Existem diversas diretrizes e manuais atualizados com orientações para a realização de revisões sistemáticas, como o Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2022), o JBI Manual for Evidence Synthesis (Aromataris; Munn, 2020), o Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care (Centre for Reviews and Dissemination, 2009), o Campbell Systematic Reviews: Policies and Guidelines (The Campbell Collaboration, 2019), entre outros. A conhecida declaração PRISMA (Preferred Reporting Items of Systematic Reviews and Meta-Analyses, versão 2009 e versão 2020) apresenta orientações e exemplos detalhados sobre como reportar uma RS, de modo a facilitar o processo de divulgação da RS de forma transparente e completa (Page et al., 2021).

Essas diretrizes, livros, guias e declarações permitem aos pesquisadores uma conduta dentro dos padrões internacionais esperados para uma RS de qualidade. Entretanto, mesmo com um amplo detalhamento de cada processo, verificam-se ainda muitas revisões autointituladas como sistemáticas que não seguem parte das diretrizes propostas e não reportam adequadamente todos os itens esperados (Ramasamy, 2022; Siddaway et al., 2019; Steil et al., 2022). Essa falha em seguir as diretrizes internacionais para RS tende a gerar resultados tendenciosos, não reproduzíveis e pouco transparentes, consequentemente sínteses de conhecimento com baixa validade científica e prática.

Por exemplo, uma recente revisão de escopo identificou, a partir de amostra de 305 RS em psicologia, muitas discrepâncias entre o que é reportado pelas RS e o que é recomendado pelo PRISMA. Um dos itens que não foi seguido na íntegra é o que se refere à seleção e ao relato das fontes de informação utilizadas para a identificação e recuperação dos estudos primários a serem sintetizados na RS (Steil et al., 2022). A falta de um detalhamento mais específico para a etapa de identificação de fontes de informação é, potencialmente, um dos motivos para a carência de padronização na escolha e na forma de relatar as fontes em RS. A partir dos achados dessa revisão, o objetivo deste artigo é explicitar como as fontes de informação têm sido reportadas em RS em psicologia e apresentar soluções práticas para as lacunas e incoerências identificadas.

As soluções práticas indicadas neste artigo foram elaboradas a partir de orientações metodológicas das instituições referência no assunto. Levamos em consideração os diferentes níveis de habilidade dos pesquisadores no uso dos recursos digitais relacionados às fontes de informação, de modo ao artigo ser útil ao maior número de pesquisadores possível. Embora os dados analisados sejam da área da Psicologia, destacamos que as recomendações aqui apresentadas são aplicáveis a todas as áreas do conhecimento, uma vez que não há uma diferenciação ou segmentação de recomendações, de forma específica, sobre o uso de fontes de informação para a realização de RS em função da área do conhecimento.

## **2 BASES DE DADOS BIBLIOGRÁFICAS, SERVIÇOS DE BUSCA E DEMAIS FONTES DE INFORMAÇÃO**

O desenvolvimento de uma RS é integralmente dependente do uso de fontes de informação. São essas fontes que permitem a recuperação dos documentos necessários para o desenvolvimento das sínteses de evidências dos achados das pesquisas. As bases de dados bibliográficas são a fonte preferencial para o processo de busca da revisão, mas há também a possibilidade da inclusão de outras fontes de informação não acadêmicas.

Um dos pontos fortes das bases de dados bibliográficas é a existência de sistemas de buscas que possibilitam múltiplas formas de recuperação dos documentos. Toda base de dados possui um sistema de busca, mas nem todo sistema de busca é uma base de dados. Essa afirmação tem o objetivo de fazer uma distinção operacional do enquadramento dos tipos de fontes de informação para RS.

Bases de dados bibliográficas podem ser definidas como uma coleção de documentos eletrônicos ou digitalizados, organizados a partir de metadados, selecionados por meio de critérios claros, formalizados em uma política de indexação, e que podem ser recuperados de múltiplas formas por meio do uso de um sistema de busca. O acervo desse tipo de fonte de informação é organizado a fim de disponibilizar artigos de periódicos científicos, anais de congressos, patentes, livros e capítulos de livros, publicações jurídicas, entre outros documentos acadêmicos e científicos (Guinchat; Menou, 1994). Nesta definição estão excluídos os repositórios institucionais, as bibliotecas digitais e virtuais, os motores de buscas, os sites institucionais e os bancos de dados, que são fontes de informação, mas não são necessariamente bases de dados bibliográficas.

O Quadro 1 classifica, de forma simplificada, como as fontes de informação acadêmicas bibliográficas podem ser divididas considerando como critérios o formato em que o conteúdo é disponibilizado (se referencial ou texto completo), o tipo de acesso à fonte de informação (se restrito ou gratuito) e o tipo de acesso aos documentos no acervo da fonte de informação (se pago, aberto ou híbrido). Uma mesma fonte de informação pode ser classificada em mais de uma categoria.

**Quadro 1** - Classificação das fontes de informação acadêmicas bibliográficas

CATEGORIAS		DESCRIÇÃO	EXEMPLOS
FORMATO DO CONTEÚDO	Referencial	Disponibilizam o acesso ao registro ou metadados do documento como título, resumo, autores, periódico, ano, palavras-chave, entre outros itens que estão indexados em uma fonte de informação. Geralmente possuem um link que direciona para uma fonte externa como uma base de dados de texto completo, para o periódico detentor do documento ou para um repositório institucional.	ERIC Embase PubMed PsycINFO Scopus Web of Science
	Texto Completo	Disponibiliza o acesso ao documento em si de forma integral para leitura e/ou <i>download</i> . A disponibilização pode ser feita no formato original do documento proveniente de periódicos científicos, como PDF ou EPUD, ou pode estar disponível no formato HTML, ou seja, neste caso o conteúdo é disponibilizado no próprio navegador web.	Annual Reviews Emerald JSTOR Science Direct SciELO
ACESSO À FONTE DE INFORMAÇÃO	Acesso Restrito	Restringem seu acesso aos usuários que assinam o serviço. Também é possível ter acesso via contratos de universidades, bibliotecas ou hospitais, por exemplo, que assinam o serviço para uso local. Há a possibilidade de acesso remoto por meio de contratos mais amplos, a exemplo, do Brasil, com o Portal de Periódicos da Capes, que disponibiliza o acesso a bases de dados pagas para indivíduos vinculados a determinados órgãos de ensino e pesquisa. Embora o acesso à base de dados ou fonte de informação seja pago, seu conteúdo pode disponibilizar documentos de acesso aberto e de acesso pago.	Academic Search Premier CINAHL Embase Engineering Village Scopus Web of Science
	Acesso Gratuito	Permitem o acesso grátis ao sistema de busca e/ou ao seu acervo, sem necessidade de assinaturas ou investimento financeiro por parte do usuário ou de instituições. Embora o acesso à base de dados seja sem custo, o seu conteúdo nem sempre é, obrigatoriamente, de acesso aberto. Em alguns casos os documentos recuperados podem ser de acesso pago, sendo necessário investimento financeiro para aquisição do documento ou a utilização de uma plataforma intermediária, como o Portal de Periódicos Capes. O acesso à base de dados não dá, necessariamente, direito ao seu acervo na totalidade.	Cochrane Dimensions Lilacs PubMed DOAJ ERIC SciELO Repositórios institucionais

ACESSO AO ACERVO	<b>Acesso Pago</b>	Os documentos disponíveis no acervo pertencem a uma editora científica, sendo necessário investimento financeiro para sua aquisição e uso. Embora o acervo seja majoritariamente de acesso pago há iniciativas para disponibilização de parte do acervo em acesso aberto.	ACM Digital Library Annual Reviews Emerald JSTOR Science Direct
	<b>Acesso Aberto</b>	Os documentos disponíveis no acervo estão sob licenças de acesso aberto, como a <i>Creative Commons</i> ou similares, sem necessidade de investimento financeiro para aquisição do documento. Um mesmo documento pode ser encontrado em diferentes fontes, pois a licença permite o compartilhamento do documento.	DOAJ SciELO Lilacs Repositórios institucionais
	<b>Acesso Híbrido</b>	O acervo é composto por documentos de acesso pago e de acesso aberto. É um modelo adotado por diversas bases de dados de acesso gratuito e que tem ganhado força em bases de dados bibliográficas de acesso restrito devido ao movimento <i>open access</i> .	Cochrane Embase PubMed Scopus

**Elaborado:** Pelos autores

As fontes de informações bibliográficas utilizadas em RS são majoritariamente acadêmicas. No entanto, a depender dos objetivos e estruturação da revisão, há possibilidade do uso de fontes não acadêmicas, como os sites governamentais, repositórios de dados estatísticos, acervos de relatórios institucionais, documentos de fontes organizacionais e até mesmo consulta a especialistas.

Todos os manuais citados na introdução deste trabalho possuem itens dedicados a recomendações para o uso de fontes de informações em RS. No entanto, devido à complexidade e constante atualização dos mecanismos e funcionalidades de busca, tais recomendações são genéricas e, em alguns casos, podem até ser limitantes para buscas mais criativas. Mesmo em um cenário de múltiplas opções de fontes de informação e de diferentes formas de executar a busca, é importante que o relato desses dois itens seja feito de forma clara e consistente. Para tal, o uso do PRISMA 2020 (Page et al. 2021) pode ser um importante mecanismo para uma descrição compreensiva e replicável do uso das fontes de informação.

### 3 PROCEDIMENTOS METODOLÓGICOS

Os dados sobre a utilização de fontes de informação em RS utilizados neste artigo advém de uma recente revisão de escopo sobre como são reportadas RS em Psicologia (Steil et al., 2022). A referida revisão avaliou o quanto as revisões autodeclaradas como sistemáticas em Psicologia seguem as recomendações do checklist PRISMA.

A busca foi realizada nas bases de dados PsycINFO e MEDLINE (via PubMed). PsycINFO foi escolhida por ser específica para a área de Psicologia e MEDLINE em função da ampla gama de artigos na área da Saúde. Uma bibliotecária auxiliou no desenvolvimento da estratégia de busca para cada base, que incluiu como cabeçalhos de assunto centrais os termos *psychology AND systematic review OR meta-analysis* em conjunto com termos livres, para revisões publicadas entre 01/01/2019 e 09/11/2020 (data da busca), sem restrição de linguagem. Todos os cabeçalhos de assunto foram explodidos e foram recuperados 2.487 RS.

Quatro autoras independentes examinaram os estudos, em um processo duplo-cego, lendo os títulos e resumos, utilizando o software Rayyan® (Ouzzani et al., 2016). Considerando o grande número de artigos que atendeu aos critérios de elegibilidade na fase de triagem (N=1472), foi realizado um cálculo amostral na plataforma OpenEpi (disponível em: <https://www.openepi.com/SampleSize/SSPropor.htm>),

considerando um intervalo de confiança de 95% e uma margem de erro de 5%. Esse procedimento originou uma amostra de 305 RS a serem incluídas na revisão. Um gerador de números aleatórios foi usado para a seleção dos artigos específicos do montante de 1472 inicialmente triados, os quais seguiram para a etapa de análise do texto completo. Quatro autoras também leram, independentemente, os textos completos, em processo duplo-cego, e as divergências foram resolvidas em reunião até que o consenso fosse alcançado. A extração dos dados das 305 RS foi realizada em uma planilha no Excel® com base nos itens do PRISMA. Para todos os dados encontrados, estatísticas descritivas foram calculadas.

A partir da triagem dos dados foram criadas planilhas em Excel® com a descrição detalhada dos itens de cada RS de acordo com os itens do PRISMA. Especificamente para este trabalho os autores analisaram os dados referentes ao uso das fontes de informação.

## **4 RESULTADOS E DISCUSSÕES**

A partir da busca e extração de dados originais de Steil et al. (2022), o presente estudo tem como foco as informações brutas extraídas acerca das fontes de informação, como relatadas pelos autores das RS. Os resultados apresentados e discutidos estão organizados em categorias.

### **4.1 Nomenclaturas das bases de dados bibliográficos**

Nossa análise dos dados sugere que, a depender da fonte de informação utilizada, existe certa dificuldade em nomear onde a pesquisa foi executada. Tomemos como exemplo a PubMed, que comumente é referenciada como uma base de dados. Embora essa ideia seja comum para alguns pesquisadores, a National Library of Medicine, proprietária da PubMed, a define como um serviço de busca que engloba três subconjuntos de dados: MEDLINE, PubMed Central e NCBI BookShelf. Ou seja, ao indicar a PubMed como uma base de dados o pesquisador está incorrendo em um erro prático, conceitual e de nomenclatura. Na verdade, a base de dados é a MEDLINE, pois é ela a responsável pela política de indexação, característica inerente à uma base de dados bibliográfica, via o Literature Selection Technical Review Committee (National Library of Medicine, 2020). No entanto, a MEDLINE não possui um sistema de busca próprio, sendo necessário utilizar serviços de busca intermediários como a PubMed, Livivo, Embase, Cochrane, Ovid, EBSCO Host, Biblioteca Virtual em Saúde (BVS), entre outras, por exemplo. Os serviços de busca se resumem a uma interface que disponibiliza um mecanismo de recuperação da informação de um acervo já selecionado, organizado e indexado. Diferentemente de uma base de dados bibliográfica, os serviços de busca, geralmente, não fazem parte e não exercem influência no processo editorial necessário para a formação do acervo, disponível nas fontes de informação, que eles disponibilizam ao usuário.

Os dados coletados na revisão de escopo (Steil et al., 2022) demonstram que há o uso de nomenclaturas não padronizadas para uma mesma fonte de informação. Para exemplificar esse ponto vamos usar o caso da base de dados MEDLINE. A revisão de escopo indicou que essa base foi nomeada como MEDLINE sem a indicação do serviço de busca (n=82), MEDLINE via PubMed (n=18), MEDLINE via Ovid (n=01), PubMed sem a indicação do subconjunto de dados da MEDLINE (n=70) e MEDLINE e PubMed simultaneamente (n=19) (Steil et al., 2022). Desse modo, em muitas RS em psicologia não fica claro, por conta da nomenclatura utilizada, qual o serviço de busca utilizado pelos pesquisadores para acessarem o acervo da MEDLINE.

Apesar de termos utilizado a MEDLINE como exemplo, outras bases de dados

bibliográficas estão sujeitas a esse mesmo problema de padronização, pois podem ser acessadas por múltiplos serviços de busca. Como exemplo podemos citar a Educational Resources Information Center (ERIC), o Food Science and Technology Abstracts, a GEOBASE, a PsycArticles, entre outras.

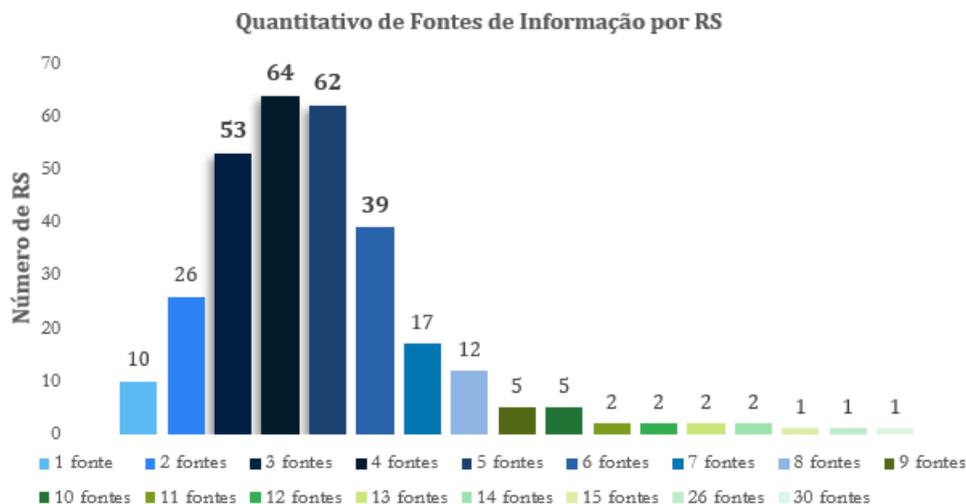
Para além de uma melhor compreensão e replicabilidade do processo de busca, a indicação correta da base de dados é um importante fator para a análise do processo de busca. Esse ponto é importante porque um mesmo conjunto de dados, acessado por meio de diferentes serviços de busca, tende a recuperar diferentes documentos, tanto em quantidade como em qualidade (Katchamart et al., 2011) em parte devido ao processo de representação temática, por exemplo. Portanto, a indicação correta da nomenclatura da base de dados auxilia o pesquisador a compreender e replicar o processo de busca de acordo com a base de dados bibliográfica e/ou serviço de busca utilizado.

Nossa sugestão é a de que seja indicado o serviço de busca intermediador utilizado para acessar uma base de dados, quando for o caso, de modo que seja possível replicar a busca de forma fidedigna. Uma forma de descrever a base de dados e o serviço de busca intermediador é, por exemplo, indicando ambos: MEDLINE via PubMed, se a restrição para esse subconjunto de dados tiver sido feita. Em serviços de busca que possibilitam a seleção da MEDLINE, mas que também buscam em outras bases de dados, é importante fazer essa indicação. Um exemplo é a EBSCO. Nesse caso, ao invés de informar que a busca foi feita na EBSCO, que é um serviço de busca único para diversas bases de dados (o EBSCO Host), é interessante indicar como: MEDLINE via EBSCO, quando for o caso. Essa orientação vale tanto para pesquisadores que utilizam a MEDLINE por meio de qualquer outro serviço de busca como para outros exemplos de bases de dados bibliográficas existentes nas diversas áreas do conhecimento.

## 4.2 Quantitativo das bases de dados

Embora não exista um consenso sobre quantas bases de dados bibliográficas são necessárias para a execução de uma revisão sistemática (Bramer et al., 2017), há a compreensão de que um maior quantitativo de fontes de informação relevantes pode reduzir possíveis riscos de viés no processo de coleta dos dados (Higgins et al., 2021). Os dados da revisão de Steil et al. (2022) chamaram atenção para a ampla diferença quantitativa de bases de dados bibliográficas utilizadas: os números variaram de 1 até 30 bases de dados em uma única RS. Não pretendemos determinar um número específico de bases de dados a serem utilizadas em uma RS, mas a partir dos dados coletados verificou-se predomínio de uso de 3 a 6 bases de dados nas RS da área de psicologia.

O Gráfico 1 apresenta o quantitativo de base de dados utilizado nos estudos incluídos na amostra.

**Gráfico 1** - Quantitativo de bases de dados bibliográficas utilizadas em RS na área de Psicologia

**Fonte:** Elaborado pelos autores com base em Steil et al. (2022).

As RS utilizaram predominantemente 3 ( $n=53$ ), 4 ( $n=64$ ), 5 ( $n=62$ ) ou 6 ( $n=39$ ) fontes de informação para executar a busca dos documentos (Steil et al., 2022). Consideramos que utilizar somente 3 bases de dados pode não abranger diferentes conjuntos de dados de pesquisa importantes para responder à pergunta da RS. Nossa sugestão é a de que, quando possível, as bases de dados escolhidas sejam de pelo menos dois tipos: a) multidisciplinares e b) específicas da área de conhecimento da pesquisa, com o adicional de uma base de dados regional. Vamos tomar como exemplo a área da educação. Em um estudo de RS os autores poderiam optar pela Scopus e Web of Science como bases de dados bibliográficas do tipo multidisciplinar, em conjunto com a Educational Resources Information Center (ERIC), Education Research Complete via EBSCO, Education Database via ProQuest e Educ@ como bases de dados bibliográficas específicas da área da educação, sendo a última uma do tipo regional, assim totalizando 6 bases de dados bibliográficas.

Entendemos que, a depender da pergunta da RS, do tipo de acesso que o pesquisador possui e do conteúdo das bases de dados bibliográficas ou fontes de informação, em conjunto com as habilidades de busca dos pesquisadores, é possível que essa sugestão não seja totalmente exequível. Nesses casos sugerimos buscar o auxílio de um profissional da área de informação, preferencialmente um bibliotecário.

### 4.3 As estratégias de busca

Estratégia de busca, também conhecida como string ou chave de busca, é a representação da pergunta de pesquisa, dos objetivos ou das hipóteses de um estudo em uma linguagem compreensível pelo sistema de busca da base de dados. O pesquisador deve traduzir sua questão em um conjunto de termos autorizados, conhecidos como cabeçalhos de assunto, que podem ser combinados com operadores booleanos, caracteres curingas, etiquetas de identificação de metadados, entre outros mecanismos, a fim de recuperar os documentos, disponíveis no acervo da base de dados, que contêm os termos definidos na estratégia de busca (Macfarlane; Russell-Rose; Shokraneh, 2022). Cabeçalhos de assunto são os termos que compõem um vocabulário controlado de um domínio de conhecimento, como um tesouro, por exemplo, e são previamente definidos por uma equipe técnica multidisciplinar. Eles têm por finalidade possibilitar o processo de representação temática de um documento em um acervo, ou seja, descrever os assuntos que aquele documento aborda, e auxiliam o usuário no

processo de recuperação desse documento quando da busca por assunto (Lazarinis, 2014).

Os dados da revisão de escopo (Steil et al., 2022) mostram que 95,73% (n=292) das RS em Psicologia informaram o uso de uma estratégia de busca para recuperação dos documentos. Deste total, 73,97% (n=216) apresentaram a estratégia de busca em seu formato de uso na base de dados, seja no corpo do texto, seja em material suplementar. Os 26,03% restantes (n=76) apresentaram somente os termos utilizados, sem apresentar a estrutura completa da estratégia de busca, impossibilitando sua replicação. Considerando a necessidade de replicabilidade nas RS, isso seria possível apenas em cerca de 70,81 % do total de revisões analisadas.

É importante compreender que uma estratégia de busca bem elaborada tende a reduzir o tempo da busca, recuperar documentos de maior utilidade e permitir a localização de ideias implícitas da pergunta de pesquisa. Quanto melhor for a estratégia de busca, melhor será o aproveitamento dos recursos da base de dados bibliográfica ou da fonte de informação. Tal importância é declarada no estudo de Salvador-Oliván, Marco-Cuenca e Arquero-Avilés (2019), onde os autores verificaram que 92,7% das RS que estão indexadas na MEDLINE via PubMed possuem alguma falha em suas estratégias de busca.

Embora haja a recomendação da apresentação das estratégias completas utilizadas em todas as bases, por exemplo, no PRISMA 2020 (Page et al., 2021) e no MECIR (Higgins et al., 2021), alguns periódicos solicitam apenas a apresentação de uma estratégia (geralmente a utilizada na MEDLINE, no caso das RS em Ciências da Saúde). Nossa recomendação é a de que os autores apresentem todas as estratégias de buscas, nos formatos utilizados em cada base de dados. Essa disponibilização tende a melhorar a replicação da recuperação da informação nos serviços de busca das bases de dados porque: a) auxilia na construção de estratégias de buscas mais eficazes por pesquisadores menos experientes, b) tem poder para padronizar determinadas buscas a partir do uso de estratégias já validadas e c) otimiza o tempo do pesquisador na elaboração da estratégia de busca e no uso eficaz dos sistemas de busca das bases de dados bibliográficas e demais fontes de informação.

Chamamos a atenção de que a replicabilidade da estratégia de busca é possível apenas quando ela é realizada na mesma base de dados na qual a estratégia de busca original foi aplicada. Isso ocorre porque diferentes bases de dados bibliográficas possuem mecanismos de buscas diferentes, assim como utilizam vocabulários distintos. Para exemplificar os vocabulários distintos, tomemos como exemplo uma RS que queira recuperar estudos com mulheres. Considerando a utilização dos cabeçalhos de assunto do tesauro específico para cada base de dados, temos: a) na PubMed, o cabeçalho de assunto para o termo mulher é “women”, b) na Embase é “female”, e c) na PsycInfo é “human females”.

Ao simularmos uma busca na Embase com a utilização do cabeçalho de assunto da PsycInfo (human females), teríamos um total de 533 documentos. Entretanto, se utilizarmos o cabeçalho de assunto correto para essa base (female), retirado do Emtree, esse número sobe para 11.079.902 documentos. Nesse sentido, nossas sugestões, para um melhor aproveitamento das bases de dados, quanto ao uso de estratégias de busca, são:

- Selecione o campo ou etiqueta correspondente à busca por cabeçalho de assunto, de acordo com o vocabulário padrão da base de dados, caso exista, a fim de recuperar os documentos indexados por aquele termo. Por exemplo, (a) na PubMed você deve utilizar o termo mais a etiqueta [MeSH Terms] (ou sua abreviação [MH]) da seguinte forma: women [MeSH Terms]; (b) em bases que possuem um campo de busca individualizado, como a Embase, Cinahl ou PsycInfo, por exemplo, selecione o campo de forma individual e inclua o cabeçalho de assunto. Para cada base de dados deve-se fazer esse processo individualmente, respeitando-se o vocabulário padrão.

- Na busca em outros campos, como título, resumo, palavras-chave do autor, por exemplo, é interessante fazer diferentes combinações de termos em linguagem natural e linguagem controlada (vocabulário padrão da base de dados). No entanto, essa combinação deve respeitar os sentidos semânticos, relacionais e lógicos a fim de manter a consistência da especificidade de busca. A combinação de diferentes termos e vocabulários ajuda a sensibilizar a busca quando se considera diferentes campos de busca.
- Em bases de dados e fontes de informação que não possuem vocabulários padrão para indexação, como Web of Science, Scopus, Science Direct, Emerald Insight e Scielo, por exemplo, você pode optar por utilizar a estratégia de busca com os termos combinados de diferentes vocabulários em conjunto com termos em linguagem natural.

Nessa etapa pode-se considerar o uso do Peer Review of Electronic Search Strategies (PRESS). O PRESS é um guia prático para a estruturação de estratégias de busca com o diferencial delas passarem por de uma revisão por pares que deve ser conduzida, de forma cega e pareada, por bibliotecários e/ou outros especialistas em informação. Ele é dividido em seis seções interrelacionadas que cobrem os aspectos das melhores práticas da recuperação da informação em bases de dados: tradução da pergunta de pesquisa; operadores booleanos e de proximidade; cabeçalhos de assunto; busca por palavra; ortografia, sintaxe e número de linhas; e limites e filtros (Mcgowan et al., 2016).

#### 4.4 Vocabulário controlado

Um ponto a ser destacado no processo de escolha e uso de bases de dados é a existência de um vocabulário controlado vinculado ao seu sistema de busca. Somente fontes de informação bibliográficas que possuem vocabulários controlados para a organização do seu acervo, independentemente do tipo de vocabulário, podem, expressamente, ser consideradas uma base de dados bibliográfica. Esse é um dos critérios de diferenciação entre uma base de dados bibliográfica e os demais tipos de fonte de informação.

Vocabulários controlados são estruturas terminológicas padronizadas por meio de rígidos critérios de organização que possibilitam nivelar o processo de comunicação entre diferentes atores humanos ou entre humanos e máquinas. Essas estruturas são compostas de termos ou frases que permitem, de forma consistente, descrever terminologicamente um domínio de conhecimento. Sua função é possibilitar a representação temática de documentos, através da indexação, utilizando os termos presentes no vocabulário, e, conseqüentemente, facilitar o processo de recuperação da informação em fontes de informação, já que a busca, em parte, irá considerar os termos já pré-definidos no vocabulário, assim evitando o uso de termos que não representam, necessariamente, o assunto a ser recuperado. Exemplos de vocabulários controlados são as listas de cabeçalhos de assunto, lista de autoridades, tesouro e ontologias (Chatterjee, 2017).

Para que a recuperação da informação ocorra de forma eficiente em bases de dados bibliográficas, é necessário o uso dos termos que compõem o tesouro (ou vocabulário controlado indicado pela fonte de informação). De forma ampla, esses termos são nomeados como termos autorizados. De forma específica, os termos autorizados que compõem um tesouro são denominados de cabeçalhos de assunto ou descritores, esse último uma nomenclatura utilizada no Brasil.

Embora nossas sugestões se pautem no uso de cabeçalhos de assunto para o processo de recuperação da informação, é necessário esclarecer que nem todas as fontes de informação utilizadas em RS possuem, necessariamente, um vocabulário controlado, preferencialmente um tesouro, vinculado ao seu processo de busca. Como já exemplificamos anteriormente, Scopus, Web of Science, Science Direct, Emerald Insight

e Scielo são exemplos de fontes de informação que não possuem um tesouro padrão vinculado ao processo de busca.

A importância da existência de um tesouro vinculado se dá pela padronização de comunicação entre o sistema de recuperação da informação e o pesquisador, facilitando que a estratégia de busca corresponda, tematicamente, com o assunto que foi previamente indexado (Chatterjee, 2017; Lazarinis, 2014). De uma forma simplificada é como se os vocabulários funcionassem como uma linguagem padrão entre o humano e a máquina. Desse modo, os vocabulários controlados facilitam a comunicação entre esses dois entes e, conseqüentemente, melhoram a compreensão do sistema de busca sobre o que o pesquisador deseja encontrar.

Neste quesito, nossa sugestão é a de que, sempre que possível, se dê preferência às fontes de informação com vocabulários controlados, no entanto, compreendemos que essa sugestão é mais exequível para bases de dados bibliográficas na área da saúde. No campo multidisciplinar podemos exemplificar um conjunto de bases de dados disponíveis na EBSCO Host como Academic Search Premier, SocINDEX e Library, Information Science & Technology, por exemplo, como bases de dados bibliográficas que possuem vocabulários controlados próprios.

Como solução para melhorar o processo de recuperação em bases sem vocabulários controlados (e até naqueles que os possuem, mas há necessidade de expandir os termos da busca em texto) sugerimos o uso de tesouros e listas de termos livres, também chamados de termos em linguagem natural. Os termos livres são os sinônimos, ou variantes, que embora não sejam termos autorizados podem auxiliar na padronização terminológica e expansão dos termos de busca. No Quadro 2 apresentamos algumas possibilidades de vocabulários controlados que podem auxiliar na escolha de termos de busca.

**Quadro 2** - Vocabulários Controlados.

<b>ÁREA DO CONHECIMENTO</b>	<b>VOCABULÁRIO CONTROLADO</b>
<b>Artes e Arquitetura</b>	Thesaurus of Art & Architecture
<b>Ciência da Informação</b>	Library and Documentation Thesaurus LISA Thesaurus LISTA Thesaurus Tesouro Brasileiro de Ciência da Informação Tesouro de Ciencias de la Documentación
<b>Ciências da Saúde</b>	APA Thesaurus CINAHL Thesaurus DECS (Descritores em Ciências da Saúde) Dentistry & Oral Subjects EMTREE (Embase Tree) MeSH (Medical Subject Headings) SPORTDiscus Dicionário de Sinônimos
<b>Ciências Sociais Aplicadas</b>	Business Source Premier Thesaurus EconLit Thesaurus SocINDEX Termos do Assunto
<b>Comunicação</b>	Communication Source Thesaurus
<b>Direito</b>	Tesouro do Supremo Tribunal Federal

<b>Educação</b>	Australian Thesaurus of Education Descriptors BRASED (Thesaurus Brasileiro da Educação) ERIC Thesaurus European Education Thesaurus
<b>Engenharias e Matemática</b>	Agrovoc Engineering Village Thesaurus GeoBase Thesaurus IEEE Thesaurus MathSciNet Thesaurus
<b>Multidisciplinar</b>	Academic Search Premier Termos do Assunto Catálogo de Autoridades da Biblioteca Nacional EuroVoc (tesauro multilingue e multidisciplinar da EU) Library of Congress Subject Headings ProQuest Thesaurus Science Direct Topics (lista de termos multidisciplinar) Thesaurus da UNESCO UNBIS (United Nations Bibliographic Information System)

**Fonte:** Elaborado pelos autores.

O uso desses vocabulários, mesmo não estando diretamente vinculados às bases de dados, permite que a recuperação dos documentos da revisão siga um padrão nas diferentes fontes de informação utilizadas porque expandirá e melhorará a qualidade dos termos de busca em texto. Destacamos que, embora nossa sugestão tenha valor prático, é necessário considerar que vocabulários de outras áreas do conhecimento podem e devem ser utilizados de forma interdisciplinar e a combinação com a linguagem natural é essencial para uma recuperação da informação mais sensível.

#### 4.5 Explosão em bases de dados

Embora nossos dados não possam subsidiar uma análise mais apurada sobre a explosão em base de dados, ela é um importante fator para a escolha das bases de dados. Quando essa técnica for empregada na busca é necessário que ela seja explicitamente relatada na RS.

A busca com explosão pode ser definida como o método de recuperação da informação em bases de dados que expande a busca pelo cabeçalho de assunto principal e a amplia para todos seus termos subordinados. Por exemplo, se você pesquisar o termo Psychology na PubMed, utilizando a etiqueta para MeSH Terms, por padrão, o sistema de busca irá considerar o termo de busca, mais todos os seus termos subordinados no tesauro da base. No caso do termo Psychology, os termos subordinados na busca via PubMed serão: Cognitive Science; Cognitive Neuroscience; Economics, Behavioral; Environmental Psychology; Ethnopsychology; Forensic Psychology; Psychology, Adolescent; Psychology, Child; Psychology, Clinical; Psychology, Comparative; Psychology, Developmental; Psychosocial Functioning; Psychology, Educational; Psychology, Experimental; Psychology, Industrial; Presenteeism; Time Management; Psychology, Medical; Psycho-Oncology; Psychology, Positive; Psychology, Social; e Psychology, Sports.

Nesse exemplo, a recuperação de um único termo de busca (Psychology) foi automaticamente expandida para vinte dois outros termos subordinados que não são, necessariamente, aqueles que o pesquisador tinha como objetivo de busca principal. Esse processo de busca ampliada só terá funcionalidade para os campos de metadados de indexação e diretamente relacionados com o vocabulário controlado da base de dados bibliográfica, não valendo para outros campos de busca, como título, resumo ou palavras-chave do autor, por exemplo.

Com relação à explosão, nossa sugestão é a de que ao escolher as bases de dados para a RS, os pesquisadores verifiquem o comportamento de busca em relação à explosão de termos. O pesquisador deve analisar se a explosão ocorre na busca padrão, como é o caso da PubMed, ou se é necessário adicionar um indicativo para que ocorra a explosão, como na Cumulative Index to Nursing and Allied Health Literature (CINAHL), por exemplo. Essa verificação é importante porque, a depender do sistema de busca, será necessário: a) adicionar caracteres, como o símbolo + na CINAHL; b) selecionar o tipo de explosão, como na Embase; c) ou a edição de tags para evitar a explosão em sistemas de busca que explodem por padrão, como na PubMed.

Embora seja útil em alguns casos, quando é necessária uma maior sensibilidade na busca, a explosão pode ser um grande problema quando o termo principal possui vários termos subordinados que não possuem relação direta com a temática principal da busca. Em alguns casos, a explosão acaba impossibilitando a pesquisa tendo em vista que os resultados são quantitativamente expressivos, o que dificulta o processo de seleção de documentos. Opte pela busca com explosão quando o conjunto de termos subordinados ao cabeçalho de assunto seja totalmente relevante para sua busca. Em casos em que parte dos termos subordinados não sejam relevantes, sugerimos o uso do cabeçalho de assunto principal em conjunto com os cabeçalhos de assunto subordinados realmente pertinentes.

É importante salientar que somente em sistemas de busca e bases de dados bibliográficas que possuem um tesouro integrado ao seu mecanismo de recuperação é possível o uso dessa funcionalidade, tal como a Medline via PubMed, a Medline via EBSCO, a Embase, a CINAHL, a Cochrane, a PsycInfo e a OVID, por exemplo. Fontes de informação como Web of Science e Scopus não possibilitam a explosão, pois não possuem um tesouro integrado. Essas duas bases optam, respectivamente, por um sistema de indexação automatizado, denominado KeyWords Plus, e pela interoperabilidade de diferentes vocabulários, que são colhidos das fontes originais em que os documentos estão disponibilizados.

#### **4.6 Literatura cinzenta**

Embora a conceituação de literatura cinzenta possa carecer de uma padronização entre diferentes áreas do conhecimento, ela pode ser compreendida como o conjunto de documentos, impressos ou digitais, que foram produzidos por diferentes atores, como entes governamentais, empresariais, industriais e acadêmicos, que não estão sob controle de editoras comerciais ou que não foram submetidos à revisão por pares, mas que contêm informações úteis que complementam a análise de um estudo e/ou servem para mitigar possíveis problemas de viés científico (Higgins et al., 2021; Paez, 2017). Uma variedade de documentos pode ser classificada como literatura cinzenta, entre eles: teses e dissertações, relatórios de pesquisas, instruções e guias de comitês, relatórios governamentais, artigos de conferências e pesquisas em andamento, capítulos de livros, dados não publicados, correspondência pessoal, dados de caráter político, entre outros (Hopewell et al., 2007; Paez, 2017).

A maioria das RS em psicologia analisadas (65,57%, n=200) não incluiu nenhuma fonte de literatura cinzenta em sua busca. Somente 34,43% (n=105) dos estudos optou pelo uso da literatura cinzenta (Steil et al., 2022). Destaca-se que, embora a literatura cinzenta possa ser frequentemente indicada como uma literatura de menor qualidade ou considerada de difícil recuperação (Woods; Phillips; Dudash, 2020), essa percepção não é compartilhada pelos manuais para revisões sistemáticas. Um exemplo claro é a indicação, pelo PRISMA 2020, do uso da literatura cinzenta como fonte de informação.

Ainda quanto à qualidade dos estudos, ressalta-se que não é adequado limitar as buscas em fontes de informação para periódicos de acordo com ranqueamentos

como Qualis CAPES, Fator de Impacto, Índice H, Cite Score, entre outros. Esse procedimento limita a localização de documentos que possam conter importantes dados para a síntese de evidências. Ademais, é papel de uma RS sintetizar todas as evidências disponíveis que respondam à sua pergunta da revisão. A análise de risco de viés, ou a avaliação de qualidade metodológica, dos estudos incluídos é que irá apresentar informações sobre os vieses potenciais dos estudos individuais incluídos ou sobre a sua qualidade metodológica. A síntese da evidência acumulada, por sua vez, mostrará a certeza da evidência sintetizada na RS.

Uma recomendação adicional sobre a literatura cinzenta é que sua forma de recuperação é diferente das bases de dados, pois geralmente ela está disponível em repositórios, motores de busca e bancos de dados específicos para esse fim. Enquanto nas bases de dados utiliza-se, em sua grande maioria, tesouros como o vocabulário controlado padrão para a representação temática dos documentos, na literatura cinzenta o processo de representação não é padronizado entre as diferentes plataformas e, comumente, não se utiliza um vocabulário controlado.

Como solução para uma recuperação mais eficaz, sugerimos o uso de estratégias de buscas alternativas com formato reduzido, que representem a temática principal da pesquisa. Essa sugestão baseia-se no fato de que algumas dessas fontes possuem limites na quantidade de termos que podem ser buscados, o que impossibilita o uso da mesma estratégia utilizada em uma base de dados bibliográfica. Em alguns casos é necessário utilizar mais de uma estratégia para se extrair resultados de busca melhores.

O Google Acadêmico é uma fonte de informação, frequentemente indicada para a recuperação de literatura cinzenta, que apresenta uma dessas limitações. A caixa de busca do Google Acadêmico tem um limite de 256 caracteres, contando com espaços. Os termos de busca que ultrapassam esse limite são ignorados pelo sistema e podem apresentar resultados fora do padrão de busca ou falsos positivos (Boeker; Vach; Motschall, 2013). Do total de 105 estudos que indicaram o uso da literatura cinzenta, 39 (37,14%) RS utilizaram o Google Acadêmico como uma das fontes de literatura cinzenta (Steil et al., 2022). Tomando essa informação como base, nossa recomendação é verificar, quando possível, as instruções de busca da fonte de informação com intuito de utilizá-la de acordo com os parâmetros padrão de busca a fim de minimizar, quando possível, suas limitações.

Para conhecimento do leitor, além do Google Acadêmico, considerado um motor de busca que possibilita encontrar literatura cinzenta (além de outros materiais que não são literatura cinzenta, pois estão indexados em bases de dados, por exemplo), existem fontes específicas para a busca de literatura cinzenta, como o Open Grey, Grey Guide e Networked Digital Library of Theses and Dissertations (NDLTD), Open Access Theses and Dissertations (OATD), Biblioteca Digital de Teses e Dissertações (BDTD), entre muitos outros.

#### **4.7 Uso de fontes complementares**

Aproximadamente 67,54% (n=206) dos trabalhos analisados na revisão de escopo (Steil et al., 2022) utilizaram algum tipo de fonte complementar, como lista de referências, contato com especialistas ou estudos de revisão anteriores, por exemplo. Desse total, 35,92% (n=74) estudos indicam o uso das listas de referências como fonte complementar (Steil et al., 2022), respeitando uma das orientações para o uso de fontes complementares em RS (Higgins et al., 2021). Embora a lista de referências tenha sido o método preferencial de acesso a fontes complementares nas RS analisadas, há outras opções pouco exploradas. Uma possibilidade é o uso da ferramenta similar articles ou similar records (a depender da base de dados), utilizada somente por um estudo da revisão de escopo. Ao selecionar um artigo específico, essa ferramenta apresenta um

conjunto de sugestões de outros artigos que possuem relações temáticas com o assunto indexado no artigo originalmente selecionado, ou seja, a plataforma oferece sugestões de documentos similares.

Mesmo com o uso de estratégias de buscas bem elaboradas, é possível que determinados documentos possam não ser recuperados por elas. Isso pode se dar pela não inclusão de um termo na estratégia de busca, a possível restrição de um operador booleano ou a uma representação temática que não corresponda necessariamente ao conteúdo presente no documento. Executamos um teste com o artigo *Systematic Review and Meta-Analysis of Self-Serving Attribution Biases in the Competitive Context of Organized Sport*, com a utilização da ferramenta *similar articles* na PubMed e Embase. Obtivemos, respectivamente, 99 e 249.470 resultados similares. Diante desse cenário, impossível quantitativamente de ser analisado na íntegra, sugerimos que o pesquisador organize os documentos indicados por relevância e defina um número padrão para a leitura como fonte complementar.

Chamamos a atenção ao uso cauteloso da ferramenta *similar articles*, uma vez que ela pode: (a) indicar documentos que já estão presentes no conjunto recuperado na busca principal; (b) sugerir artigos sem relação lógica com a pergunta de pesquisa da revisão, pois o algoritmo considerou a presença de determinados cabeçalhos de assunto ou termos gerais e não necessariamente respeitou a lógica dos operadores booleanos; e (c) indicar documentos que não sejam aqueles definidos no protocolo de busca da pesquisa.

#### **4.8 Regionalidade das fontes de informação**

Os resultados da revisão de escopo (Steil et al., 2022) indicam um reduzido (ou limitado) uso de bases de dados regionais. A partir de uma avaliação adicional identificamos que, do montante da amostra de 305 estudos, somente 7,86% (n=24) utilizaram pelo menos uma fonte de informação regional.

Claramente há uma preferência por bases de dados mais consolidadas, localizadas em países de maior relevância científica. No entanto, existem iniciativas regionais que podem e devem ser consideradas como fontes de informação relevantes para a coleta de dados em RS, como *Literatura Latino-Americana e do Caribe em Ciências da Saúde* (Lilacs), *Scientific Electronic Library* (SciELO), *China National Knowledge Infrastructure*, *Africa-Wide Information*, entre outras. Uma fonte de informação regional é aquela que indexa somente conteúdo de um determinado país ou região geográfica, em um ou mais idiomas.

Tomemos como exemplo as RS publicadas por pesquisadores brasileiros e chineses presentes na revisão de escopo de Steil et al. (2022). No primeiro caso, de um total de 7 estudos, 71,42% (n=5) utilizaram pelo menos uma fonte de informação de nível regional. No segundo caso, de 19 estudos, somente 31,57% (n=6) utilizaram alguma fonte de informação regional. Não temos dados suficientes para apontar uma conclusão sólida, mas o maior uso de bases regionais em estudos com pesquisadores brasileiros, em nossa análise, pode ser um resultado direto de anos de investimento em recursos humanos, infraestrutura tecnológica, marketing e divulgação científica por parte do Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde - Bireme e da iniciativa SciELO, entre outras iniciativas regionais. Especificamente no caso chinês, podemos supor que a necessidade de publicar em periódicos estrangeiros seja uma variável de influência, pois no caso brasileiro 3 publicações foram feitas em periódicos regionais e todas utilizaram fontes de informação regionais. No caso dos pesquisadores chineses, embora tenham utilizado fontes de informação regionais, nenhum dos estudos presentes na amostra foi publicado em um periódico regional.

Nossa sugestão é a de que, sempre que possível, os pesquisadores incluam no mínimo uma base de dados regional no processo de coleta de dados,

sugestão essa já existente em outras orientações, como o MERCIR (Higgins et al., 2021), por exemplo, mas pouco presente em outras orientações internacionais.

Ao executar a busca exclusivamente em bases de dados internacionais, o pesquisador pode, não intencionalmente, excluir documentos relevantes, mas que não estão, por algum motivo, indexados nessas fontes de informação (Betrán et al., 2005). No âmbito prático, não há motivos para a exclusão das bases de dados regionais, tendo em vista que sua profissionalização quanto à infraestrutura, critérios de indexação, usabilidade e qualificação dos recursos humanos têm, cada vez mais, se equiparado às bases de dados bibliográficas internacionais.

#### **4.9 Papel do bibliotecário**

Independentemente do nível de habilidade de pesquisa da equipe da RS, é recomendada a participação de um bibliotecário a fim de auxiliar todo o processo de gestão de dados da pesquisa. Esse profissional pode estar presente em todas as fases da pesquisa a fim de possibilitar um melhor uso das ferramentas e funcionalidades das bases de dados bibliográficas, sendo essa uma recomendação de diversas diretrizes de revisão (Aromataris; Munn, 2020; Higgins et al., 2022). Embora seja uma recomendação, somente 9,83% (n=30) das RS em Psicologia tinham esse profissional em sua equipe (Steil et al., 2022).

Os estudos que tinham um bibliotecário foram mais propensos a apresentar a estratégia de busca (80%) quando comparados àqueles que não o tinham (69,81%). O número médio de bases de dados bibliográficas utilizadas foi similar (com a participação de um bibliotecário foi de 4.6 e sem um bibliotecário foi 4.9).

Embora nossos resultados não possam afirmar se a participação desse profissional exerce papel na qualidade informacional das RS, nossa sugestão é a de que, sempre que possível, ele esteja na equipe de pesquisa, pois "a contribuição de um bibliotecário pesquisador ou cientista da informação pode ser inestimável para projetar e refinar a pesquisa" (Aromataris; Munn, 2020, sem paginação).

### **5 CONCLUSÕES**

O objetivo deste artigo foi a de apresentar soluções para problemas identificados em revisões sistemáticas publicadas ao que tange o uso e relato das fontes de informação bibliográficas. A partir da nossa análise sugerimos as seguintes soluções:

- Indicar corretamente em que fonte de informação foi feita a coleta de dados, relatando o nome da base de dados bibliográficas, fonte de informação ou serviço de busca utilizado para pesquisar o conjunto de dados;
- Definir entre 3 e 6 bases de dados para a coleta de dados e, quando possível, combinar bases de dados bibliográficas multidisciplinares com bases disciplinares;
- Dar preferência para bases de dados bibliográficas que utilizam um vocabulário controlado como padrão de representação temática a fim de obter um processo de busca mais consistente e uma recuperação mais eficaz;
- Quando optar por bases de dados com vocabulário controlado integrado, verificar a disponibilidade de explosão dos termos de busca e informar se utilizou ou não a explosão no relato da busca na seção de procedimentos metodológicos da RS;
- Quando a literatura cinzenta for utilizada como fonte de informação fazer adaptações nas estratégias de busca tendo em vista que o sistema de busca dessas ferramentas é diferente daquele encontrado em bases de dados;
- Quando da utilização de fontes complementares por meio das bases de dados, sugere-se o uso da ferramenta de

- busca “documentos similares” considerando ser opção viável e rápida;
- Utilizar no mínimo 1 base de dados regional, uma vez que estudos relevantes podem estar publicados em fontes regionais, mas, por algum motivo, não serem recuperados em bases de dados bibliográficas internacionais;
  - Sempre que possível, consulte um bibliotecário especialista em recuperação da informação a fim de auxiliar no protocolo de busca e no processo de recuperação, seguindo os padrões de qualidade informacional, com intuito de aproveitar todas as funcionalidades das bases de dados bibliográficas.

Por fim, compreendemos que a complexidade metodológica para o desenvolvimento de uma RS impõe ao pesquisador uma série de desafios. Nesse sentido, esperamos que os esclarecimentos conceituais e as recomendações práticas deste artigo possam contribuir para o aperfeiçoamento das RS realizadas por pesquisadores em todas as áreas do conhecimento.

## REFERÊNCIAS

- AROMATARIS, Edoardo; MUNN, Zachary (ed.). **JBIM Manual for Evidence Synthesis**. Adelaide: JBI, 2020. DOI: <https://doi.org/10.46658/JBIMES-20-01>.
- BRAMER, Wichor M. et al. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. **Systematic Reviews**, v. 6, n. 1, p. 1-12, Dec. 2017. DOI: <http://dx.doi.org/10.1186/s13643-017-0644-y>.
- BETRÁN, Ana P. et al. Effectiveness of different databases in identifying studies for systematic reviews: experience from the WHO systematic review of maternal morbidity and mortality. **BMC Medical Research Methodology**, v. 5, n. 1, p. 1-5, 28 Jan. 2005. DOI: <http://dx.doi.org/10.1186/1471-2288-5-6>.
- BOEKER, Martin; VACH, Werner; MOTSCHALL, Edith. Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. **BMC Medical Research Methodology**, v. 13, n. 1, p. 1-12, 26 out. 2013. DOI: <http://dx.doi.org/10.1186/1471-2288-13-131>.
- BORREGO, Maura; FOSTER, Margaret J.; FROYD, Jeffrey E. What is the State of the Art of Systematic Review in Engineering Education? **Journal Of Engineering Education**, [v. 104, n. 2, p. 212-242, abr. 2015. DOI: <http://dx.doi.org/10.1002/jee.20069>.
- CENTRE FOR REVIEWS AND DISSEMINATION. **Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care**. York, England: University of York, 2009.
- CHATTERJEE, Amitabha. **Elements of Information Organization and Dissemination**. Cambridge: Chandos Publishing, 2017.
- GRAINGER, Rebecca et al. Issues in reporting of systematic review methods in health app-focused reviews: a scoping review. **Health Informatics Journal**, v. 26, n. 4, p. 2930-2945, Sep. 2020. DOI: <http://dx.doi.org/10.1177/1460458220952917>.
- GUINCHAT, Claire; MENOUE, Michel. **Introdução geral as ciências e técnicas da informação e documentação**. Tradução Míriam Vieira da Cunha. 2. ed. Brasília, DF: IBICT, 1994.
- HANSEN, Henrik; TRIFKOVIC, Neda. **Systematic Reviews: Questions, Methods and Usage**. Copenhagen, Denmark: Ministry Of Foreign Affairs Of Denmark, 2013.
- HIGGINS, Julian Pt; GREEN, Sally (ed.). **Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0**: updated March 2011. London, UK: The Cochrane Colaboration, 2011.
- HIGGINS, Julian P.T.; LASSERSON T; CHANDLER, Jackie; TOVEY D; THOMAS, James; FLEMYNG E, Churchill R. **Methodological Expectations of Cochrane Intervention Reviews**. London, UK: Cochrane, 2021.
- HIGGINS, Julian P.T.; THOMAS, James;

CHANDLER, Jackie; Cumpston M, Li T, Page MJ, Welch VA (ed.). **Cochrane Handbook for Systematic Reviews of Interventions versão 6.3** (Last updated: 4 August, 2022). Cochrane, 2022. Disponível em: <https://training.cochrane.org/handbook>. Acesso em: 30 jan. 2023.

HOPEWELL, Sally et al. Grey literature in meta-analyses of randomized trials of health care interventions. **Cochrane Database of Systematic Reviews**, p. 1-15, Apr. 2007. DOI: <http://dx.doi.org/10.1002/14651858.mr000010.pub3>.

KATCHAMART, Wanruchada et al. PubMed had a higher sensitivity than Ovid-MEDLINE in the search for systematic reviews.

**Journal of Clinical Epidemiology**, v. 64, n. 7, p. 805-807, July 2011. DOI: <http://dx.doi.org/10.1016/j.jclinepi.2010.06.004>.

LAZARINIS, Fotis. **Cataloguing and Classification: an introduction to AACR2, RDA, DDC, LCC, LCSH and MARC 21 standards**. Cambridge: Chandos Publishing, 2014.

MACFARLANE, Andrew; RUSSELL-ROSE, Tony; SHOKRANEH, Farhad. Search strategy formulation for systematic reviews: issues, challenges and opportunities. **Intelligent Systems With Applications**, v. 15, p. 1-10, Sep. 2022. DOI: <https://doi.org/10.1016/j.iswa.2022.200091>.

MCGOWAN, Jessie et al. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. **Journal of Clinical Epidemiology**, v. 75, p. 40-46, July. 2016. DOI: <http://dx.doi.org/10.1016/j.jclinepi.2016.01.021>.

MELNIK, Tamara; SOUZA, Wanderson Fernandes de; CARVALHO, Marcele Regine de. A importância da prática da psicologia baseada em evidências: aspectos conceituais, níveis de evidência, mitos e resistências. **Revista Costarricense de Psicología**, San José, Costa Rica, v. 33, n. 2, p. 79-92, jul./dez. 2014.

MURAD, M Hassan et al. New evidence pyramid. **Evidence Based Medicine**,

v. 21, n. 4, p. 125-127, June 2016. DOI: <http://dx.doi.org/10.1136/ebmed-2016-110401>.

NATIONAL LIBRARY OF MEDICINE. MEDLINE, PubMed, and PMC (PubMed Central): how are they different? 2020. Disponível em: <https://www.nlm.nih.gov/bsd/difference.html>. Acesso em: 30 jan. 2023.

NEWMAN, Mark; GOUGH, David. Systematic Reviews in Educational Research: methodology, perspectives and application. In: ZAWACKI-RICHTER, Olaf et al (ed.). **Systematic Reviews in Educational Research: methodology, perspectives and application**. Cham, Switzerland: Springer, 2020. p. 3-22.

OUZZANI, Mourad; HAMMADY, Hossam; FEDOROWICZ, Zbys; ELMAGARMID, Ahmed. Rayyan—a web and mobile app for systematic reviews. **Systematic Reviews**, v. 5, n. 1, p. 1-10, Dec. 2016. DOI: <http://dx.doi.org/10.1186/s13643-016-0384-4>.

PAEZ, Arsenio. Gray literature: an important resource in systematic reviews. **Journal of Evidence-Based Medicine**, v. 10, n. 3, p. 233-240, Aug. 2017. DOI: <http://dx.doi.org/10.1111/jebm.12266>.

PATOLE, Sanjay. Systematic Reviews, Meta-Analysis, and Evidence-Based Medicine. In: PATOLE, Sanjay (ed.). **Principles and Practice of Systematic Reviews and Meta-Analysis**. Cham, Switzerland: Springer, 2021. p. 1-10.

PARK, Ho Young et al. Quality Reporting of Systematic Review and Meta-Analysis According to PRISMA 2020 Guidelines: results from recently published papers in the Korean Journal of Radiology. **Korean Journal of Radiology**, v. 23, n. 3, p. 355-369, 2022. DOI: <http://dx.doi.org/10.3348/kjr.2021.0808>.

PAGE, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **BMJ**, v. 372, p. 1-9, Mar. 2021. DOI: <http://dx.doi.org/10.1136/bmj.n71>.

PURSSELL, Edward; MCCRAE, Niall. **How to Perform a Systematic Literature Review: a guide for healthcare researchers**,

practitioners and students. Cham, Switzerland: Springer, 2020.

RAMASAMY, Akilesh. PRISMA 2020: key changes and implementation aspects.

**Journal of Oral and Maxillofacial Surgery**, v. 80, n. 5, p. 795-797, May 2022. DOI: <http://dx.doi.org/10.1016/j.joms.2021.12.018>.

SALVADOR-OLIVÁN, José Antonio; MARCO-CUENCA, Gonzalo; ARQUERO-AVILÉS, Rosario. Errors in search strategies used in systematic reviews and their effects on information retrieval. **Journal of The Medical Library Association**, v. 107, n. 2, p. 210-221, Apr. 2019. DOI: <http://dx.doi.org/10.5195/jmla.2019.567>.

SIDDAWAY, Andy P.; WOOD, Alex M.; HEDGES, Larry V. How to do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. **Annual Review of Psychology**, v. 70, n. 1, p.747-770, Jan. 2019. DOI: <http://dx.doi.org/10.1146/annurev-psych-010418-102803>.

STEIL, Andrea Valéria et al. Reporting characteristics of systematic reviews in Psychology: a scoping review. **Journal Of Health Psychology**, v. 27, n. 13, p. 2964-2981, Jan. 2022.

THE CAMPBELL COLLABORATION. **Campbell Systematic Reviews: Policies and Guidelines**. Oslo, Norway: The Campbell Collaboration, 2019. DOI: <http://dx.doi.org/10.4073/cpg.2016.1>.

WOODS, Stephen; PHILLIPS, Kathleen; DUDASH, Andrew. Grey literature citations in top nursing journals: a bibliometric study. **Journal of the Medical Library Association**, v. 108, n. 2, p. 262-269, Apr. 2020. DOI: <http://dx.doi.org/10.5195/jmla.2020.760>.