

Além da Automação: Decifrando o Alinhamento de Aplicações de Inteligência Artificial Com os Manuais de Revisões Sistemáticas

Beyond Automation: Decoding the Alignment of Artificial Intelligence Applications with Systematic Review Guidelines

Más Allá de la Automatización: Descifrando la Alineación de las Aplicaciones de Inteligencia Artificial con las Guías de Revisiones Sistemáticas

RESUMO

Revisões Sistemáticas (RS) representam uma metodologia consolidada para a síntese de evidências científicas na área da saúde, sua condução exige rigor metodológico, preconizado pelos manuais JBI e Cochrane. Avanços tecnológicos, como a Inteligência Artificial (IA) foram integrados às RS, automatizando etapas e otimizando recursos. Este estudo identificou como as aplicações baseadas em IA utilizadas na elaboração de RS da área da saúde se alinham a estes manuais, avaliando 29 estudos que empregaram IA em diferentes etapas da RS. A análise revelou que 51,7% (15 estudos) atenderam aos manuais, enquanto os 48,3% (14 estudos) não atenderam. A etapa de Seleção (primeira triagem), representou 89,7% dos estudos (26 de 29). Enquanto etapas como formulação de estratégia de busca, avaliação de risco de viés e síntese de resultados não foram abordadas. Conclui-se que, para garantir a confiabilidade das RS apoiadas por IA, é necessário alinhar essas ferramentas às diretrizes metodológicas dos manuais, bem como de um esforço conjunto entre desenvolvedores de softwares e a comunidade científica.

DESCRIPTORIOS: Revisões sistemáticas como assunto, Metodologia – revisão sistemática, Armazenamento e recuperação da informação – saúde, Inteligência artificial, Aprendizado de máquina.

ABSTRACT

Systematic Reviews (SR) represent a well-established methodology for synthesizing scientific evidence in the healthcare field, and their conduct requires methodological rigor as outlined in the JBI and Cochrane manuals. Technological advances, such as Artificial Intelligence (AI), have been integrated into SRs, automating stages and optimizing resources. This study identified how AI-based applications used in the development of healthcare SRs align with these manuals, evaluating 29 studies that employed AI in different stages of SRs. The analysis revealed that 51.7% (15 studies) adhered to the manuals, while 48.3% (14 studies) did not. The Selection stage (first screening) represented 89.7% of the studies (26 out of 29). Stages such as search strategy formulation, risk of bias assessment, and results synthesis were not addressed. It is concluded that to ensure the reliability of AI-supported SRs, it is necessary to align these tools with the methodological guidelines of the manuals, as well as foster collaboration between software developers and the scientific community.

KEYWORDS: Systematic Reviews as Topic, Methodology – systematic review, Information Storage and Retrieval – health, Artificial Intelligence, Machine Learning.

RESUMEN

Las Revisiones Sistemáticas (RS) representan una metodología consolidada para la síntesis de evidencia científica en el área de la salud, y su realización requiere un rigor metodológico, según lo estipulado por los manuales JBI y Cochrane. Avances tecnológicos, como la Inteligencia Artificial (IA), han sido integrados a las RS, automatizando etapas y optimizando recursos. Este estudio identificó cómo las aplicaciones basadas en IA utilizadas en la elaboración de RS en el área de la salud se alinean con estos manuales, evaluando 29 estudios que emplearon IA en diferentes etapas de las RS. El análisis reveló que el 51,7% (15 estudios) cumplió con los manuales, mientras que el 48,3% (14 estudios) no lo hizo. La etapa de Selección (primer filtro) representó el 89,7% de los estudios (26 de 29). Mientras que etapas como la formulación de estrategias de búsqueda, la evaluación del riesgo de sesgo y la síntesis de resultados no fueron abordadas. Se concluye que, para garantizar la fiabilidad de las RS apoyadas por IA, es necesario alinear estas herramientas con las directrices metodológicas de los manuales, así como fomentar la colaboración entre los desarrolladores de software y la comunidad científica.

PALABRAS CLAVE: Revisiones sistemáticas como tema, Metodología – revisión sistemática, Almacenamiento y recuperación de información – salud, Inteligencia artificial, Aprendizaje automático.

RECEBIDO EM: 27/01/2025 APROVADO EM: 05/02/2025

Como citar este artigo: Zanela M, Carvalho DR, Junior RM. Além da Automação: Decifrando o Alinhamento de Aplicações de Inteligência Artificial Com os Manuais de Revisões Sistemáticas. Saúde Coletiva (Edição Brasileira) [Internet]. 2025 [acesso ano mês dia];15(93):14751-14760. Disponível em: DOI: 10.36489/saudecoletiva.2025v15i93p14751-14760

ID Mariluci Zanela
Doutoranda do Programa de Pós-graduação em Gestão da Informação da Universidade Federal do Paraná.
ORCID: <https://orcid.org/0009-0001-5436-5609>

ID Ricardo Mendes Junior
Professor do Programa de Pós-graduação em Gestão da Informação da Universidade Federal do Paraná.
ORCID: <https://orcid.org/0000-0003-4947-0364>

ID Deborah Ribeiro Carvalho
Professora do Programa de Pós-graduação em Gestão da Informação da Universidade Federal do Paraná.
ORCID: <https://orcid.org/0000-0002-9735-650X>

INTRODUÇÃO

As Revisões Sistemáticas (RS) na área da saúde são estudos que utilizam uma metodologia transparente e imparcial para identificar, avaliar e sintetizar as evidências científicas^{1,2}. Seguem métodos definidos e reprodutíveis, permitindo a consolidação de estudos previamente avaliados³.

Para assegurar a qualidade e a padronização das RS, utilizam-se diretrizes metodológicas, como o Manual Joanna Briggs Institute⁴ e o Cochrane Handbook⁵. Esses manuais fornecem orientações para a condução de RS baseadas em evidências, contemplando critérios como viabilidade, adequação, relevância e eficácia das intervenções em saúde⁶.

No entanto, a execução de uma RS é um processo intensivo e desafiador, que requer a triagem de um grande volume de estudos, o gerenciamento de dados heterogêneos e um controle rigoroso para minimizar vieses, ou seja, erros metodológicos na seleção e análise dos estudos⁷.

Ferramentas baseadas em IA já estão sendo aplicadas na elaboração de RS em etapas como a triagem inicial, em que algoritmos de reconhecimento de padrões ajudam a classificar automaticamente grandes volumes de literatura,

acelerando o processo e minimizando riscos de vies. Assim, é possível otimizar etapas críticas do processo, reduzindo a carga de trabalho dos pesquisadores, melhorando a precisão na triagem de estudos e aumentando a consistência na extração de dados⁸.

Aplicações de IA utilizadas na elaboração de RS incluem métodos como Machine Learning (aprendizado supervisionado, não supervisionado e por reforço), redes neurais profundas (Deep Learning), algoritmos como Support Vector Machine (SVM) e Naïve Bayes, além de tecnologias avançadas de Processamento de Linguagem Natural (PLN), como Bidirectional Encoder Representations from Transformers (BERT) e modelos generativos, como Generative Pre-trained Transformer (GPT)⁹.

Diante desse contexto, este estudo tem por objetivo identificar como as aplicações baseadas em IA utilizadas na elaboração de revisões sistemáticas da área de saúde se alinham aos manuais JBI e Cochrane.

A justificativa para a pesquisa está na crescente adoção de ferramentas baseadas em IA como suporte à condução de RS. Essas ferramentas, embora capazes de aprimorar etapas do processo, exigem conformidade com critérios de transparência, sistemati-

dade e reprodutibilidade. Apesar disso, até o momento, não foram encontrados estudos que avaliem, de forma sistemática, o alinhamento dessas aplicações às diretrizes dos manuais JBI e Cochrane, o que evidencia uma lacuna crítica na literatura científica. Dessa forma, este artigo busca preencher essa lacuna, contribuindo para uma compreensão dessas tecnologias no campo das RS de saúde.

MÉTODO

Esta pesquisa foi conduzida por meio de revisão bibliográfica exploratória, visando identificar, a partir de RS publicadas, informações para responder ao objetivo do estudo¹⁰. O processo metodológico seguiu as etapas de: definição de uma estratégia de busca, seleção das fontes de dados, triagem e extração das informações e, por fim, síntese e interpretação dos dados extraídos dos estudos selecionados.

A estratégia utilizada na busca foi: ('artificial intelligence' OR 'machine intelligence' OR 'machine learning software' OR 'machine learning program' OR 'data processing' OR 'natural language processing' OR 'semi supervised machine learning' OR 'artificial intelligence chatbot' OR 'artificial intelligence software') AND ('systematic review (topic)' OR 'systematic reviews'

OR 'study selection' OR 'screening' OR 'screening method' OR 'citation analysis' OR 'methodological studies' OR 'methodological problems' OR 'methodological quality' OR 'method detection limit' OR 'qualitative research' OR 'qualitative study' OR 'search algorithm' OR 'search strategy').

As bases de dados selecionadas foram: MEDLINE/PubMed (National Library of Medicine), Embase (Elsevier), BVS/LILACS (BIREME), CINAHL (EBSCOhost), Cochrane Library, Scopus (Elsevier), Web of Science (Clarivate Analytics), Preprints e OpenGrey.

Foram selecionados artigos de RS que relatam o uso de aplicações de IA na elaboração de RS na área de saúde. Neste estudo, as ferramentas foram definidas como softwares, plataformas, frameworks, APIs, Chatbots e assistentes virtuais, métodos, heurísticas ou modelos de IA pré-treinados, automáticos e/ou semiautomáticos. Foram excluídos estudos que utilizavam softwares ou métodos baseados em IA apenas para o gerenciamento de dados ou para auxiliar no diagnóstico e/ou tratamento de doenças, bem como outros tipos de documentos médicos, como prontuários eletrônicos.

Os artigos coletados nas bases de dados foram armazenados no software Rayyan. Nele, realizou-se a retirada de referências duplicadas, a seleção dos estudos incluídos nesta revisão, a primeira triagem e o registro das exclusões.

Para a extração, discussão e apresentação de resultados, utilizou-se uma planilha estruturada no Microsoft Excel®. Nessa planilha, foram registrados aspectos como autor, ano de publicação, aplicação de IA e etapas da RS, em conformidade com os Manuais JBI e Cochrane.

A categorização das etapas da RS também seguiu as diretrizes metodológicas dos manuais JBI e Cochrane, que apresentam orientações similares, conforme demonstrado no Quadro 1.

Quadro 1 - Etapas da Revisão Sistemática	
Etapa	Descrição
PLANEJAMENTO	
Formulação da Pergunta da RS	Definir uma pergunta clara e objetiva utilizando acrônimos de pesquisa (PICO, PICOT, PECOS)
Critérios de elegibilidade	Definir critérios de inclusão e exclusão para determinar quais textos serão incluídos na revisão. Esses critérios devem ser baseados na questão de pesquisa e devem considerar fatores como tipo de estudo, população, intervenção, desfecho e período de publicação.
Estratégia de busca	A estratégia de busca deve ser planejada utilizando descritores padronizados da saúde como o MeSH e operadores booleanos para maximizar a sensibilidade e a especificidade da busca.
Bases de Dados	A escolha das bases de dados deve estar relacionada à pergunta de pesquisa, podendo ser bases de saúde gerais como MedLine, Embase, Lilacs, além das bases específicas como Cinahl, Cochrane Library, PEDro, PsycInfo, além das bases multidisciplinares como Scopus, Web of Science e literatura cinzenta: Google Scholar, Open Grey
Protocolo	Elaborar e registrar um protocolo detalhado e transparente, descrevendo os métodos e estratégias a serem utilizados, registrar o protocolo em uma das plataformas: PROSPERO, Cochrane Database of Systematics Reviews, JBI Evidence Synthesis
EXECUÇÃO	
Localização dos estudos	Realizar as buscas nas bases de maneira sistemática, simultaneamente, e exportar os resultados para gerenciadores de referências
Retirada de registros duplicados	Retirada de referências duplicadas através de gerenciadores de referências
Seleção de estudos	A seleção dos estudos deve ocorrer em duas triagens: 1ª triagem: leitura de títulos e resumos; por 2 revisores as cegas. 2ª triagem: leitura do texto completo: por 2 revisores as cegas. Ambas conduzidas por revisores independentes e, em caso de discordância, um terceiro revisor é acionado, assegurando menor influência de vieses. Os trabalhos elegíveis após a triagem inicial são analisados em profundidade, sendo recomendado justificar a exclusão de cada estudo e manter um registro dessas decisões através de um fluxograma (PRISMA, 2020).
Avaliação da qualidade	Estabelecer a qualidade metodológica, a confiança dos estudos, avaliação do risco de viés através de instrumentos (Rob2, Robin-s; JBI Sumari)

Revisão de Literatura

Zanela M, Carvalho DR, Junior RM

Além da Automação: Decifrando o Alinhamento de Aplicações de Inteligência Artificial Com os Manuais de Revisões Sistemáticas

Extração dos dados	Definir previamente as informações que serão extraídas dos artigos como: tipo de estudo, local e período em que foi realizado, critérios de inclusão e exclusão, tempo de seguimento, número de participantes, intervenção, desfecho apresentado (tabela de extração). Essa etapa deve ser realizada em pares de forma independente
Análise estatística	É recomendada quando os dados de mais de um estudo forem homogêneos, adequados em relação à metodologia empregada e com seus respectivos desfechos clínicos semelhantes.
Síntese dos resultados	Apresentar uma análise qualitativa ou quantitativa (síntese pode ser narrativa, metanálise, gráfica ou tabular)
Avaliação da Qualidade da evidência	Avaliar a certeza da evidência para cada desfecho analisado na revisão utilizando o sistema Grading of Recommendations Assessment, Development and Evaluation - GRADE
RELATÓRIO	
Relatório	Divulgar os achados da RS de maneira transparente, seguindo diretrizes do Guia de relato PRISMA 2020

Fonte: Os autores, adaptado de JBI⁴, Cochrane⁵.

Para avaliar o alinhamento com as recomendações dos manuais JBI e Cochrane nas revisões sistemáticas que utilizaram aplicações de IA, foram estabelecidos os seguintes critérios:

Atende: a aplicação de IA realiza etapas da RS em alinhamento com as os manuais JBI e Cochrane (Quadro 1);

Não Atende: a aplicação de IA realiza etapas da RS em desacordo com os manuais JBI e Cochrane (Quadro 1).

RESULTADOS

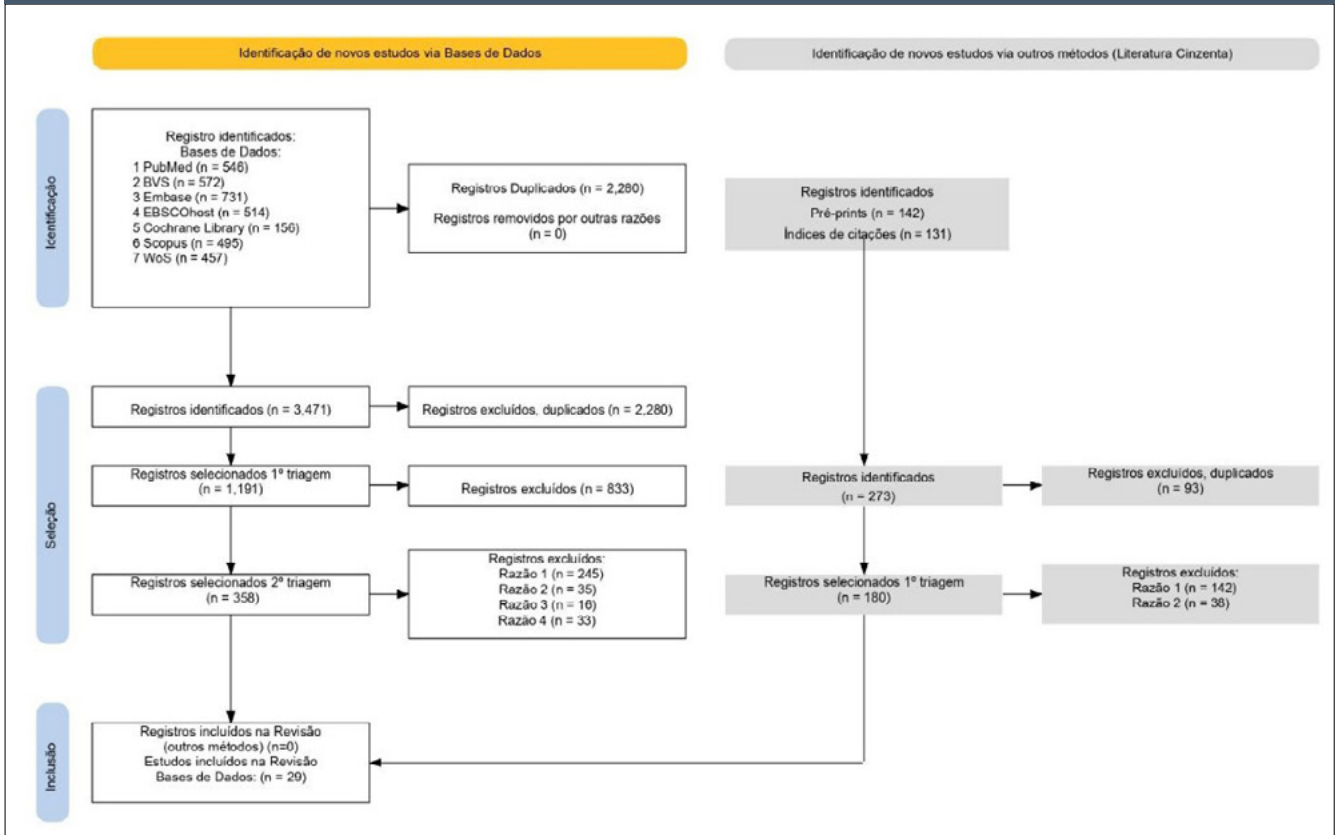
As buscas foram realizadas simultaneamente, no dia 22 de junho de

2024, nas bases de dados e resultaram em 3.471 estudos. Destes, 2.280 foram excluídos, por serem referências duplicadas, restando 1.191 para a 1ª triagem. Nessa fase, foram excluídos 833 estudos que não tinham relação com o objetivo da revisão. Restaram 358 artigos para a 2ª triagem. Nesse momento, foram excluídos 309 trabalhos, pelos motivos: 245 estudos em que a IA não era usada para elaboração de revisões (razão 1); 35 estudos excluídos por não usarem ferramentas ou métodos de IA (razão 2); 16 estudos excluídos por serem estudos de revisão ou estudos secundários (razão 3); 33 estudos excluídos de Ensaio Clínico Randomizados, ou estudos de

intervenção que não faziam parte dessa revisão (razão 4). Assim, chegou-se ao total de 29 trabalhos selecionados para extração de dados.

Além disso, as buscas na literatura cinzenta resultaram em 273 estudos. Após a remoção de 93 duplicados, restaram 180 estudos para a 1ª triagem. Destes, 180 foram excluídos pelos motivos: 142 pré-prints (razão 1) e 38 não atendiam aos critérios de elegibilidade desta revisão (razão 2). Nenhum estudo da literatura cinzenta foi incluído neste artigo (Figura 1).

Figura 1 - Fluxograma PRISMA - identificação dos estudos



FONTE: Os autores, adaptado de Page et al.¹.

Foram analisados 29 estudos publicados entre 2012 e 2023. Esses estudos avaliaram aplicações de IA em diferentes etapas da elaboração de RS

na área de saúde. Para identificar o alinhamento das aplicações de IA com os manuais JBI e Cochrane, os resultados foram organizados conforme o Quadro 1 – Etapas da RS. Cada aplicação foi

categorizada como "Atende" ou "Não Atende", conforme apresentado no Quadro 2:

Quadro 2 - Estudos incluídos na revisão			
Autores / Ano	Aplicações de IA	Etapas da RS	Alinhamento com os Manuais JBI Cochrane
Bekhuis, Demner-Fushman, 2012 ¹¹	SVM; K-NN; NB; CNB; EvoSVM	Seleção 1ª triagem	Não Atende
Jonnalagadda, Petitti, 2013 ¹²	NB; SVM; FCNB	Seleção 1ª triagem	Não Atende
Kim, Choi, 2014 ¹³	SVM	Seleção 1ª triagem	Não Atende
Blake, Lucic, 2015 ¹⁴	GLM; SVM	Seleção 1ª triagem	Não Atende
Rathbone; Hoffmann; Glasziou, 2015 ¹⁵	Abstrackr	Seleção 1ª triagem	Atende
Hashimoto, Kontonatsios, Miwa, Ananiadou, 2016 ¹⁶	Rede Neural	Seleção 1ª triagem	Não Atende
Przybyła, Brockmeier, Kontonatsios, et al. 2018 ¹⁷	Robot Analyst	Seleção 2ª triagem	Atende

Revisão de Literatura

Zanela M, Carvalho DR, Junior RM

Além da Automação: Decifrando o Alinhamento de Aplicações de Inteligência Artificial Com os Manuais de Revisões Sistemáticas

Tsafnat, Glasziou, Karystianis, Coiera, 2018 ¹⁸	GATE	Seleção 1ª triagem	Atende
Bucheli Guerrero, 2019 ¹⁹	SASR	Seleção 1ª triagem	Não Atende
Gartlehner, Wagner, Lux, et al., 2019 ²⁰	DistillerAI	Seleção 1ª triagem	Atende
Gates, Guitard, Pillay, et al. 2019 ²¹	Abstrackr; DistillerAI RobotAnalyst	Seleção 1ª triagem	Atende
Gates, Gates, Sebastianski, et al. 2020 ²²	Abstrackr	Seleção 1ª triagem	Atende
Howard, Phillips, Tandon, et al. 2020 ²³	SWIFT-Active Screener	Seleção 1ª triagem	Atende
Orgeolet, Foulquier, Misery, et al. 2020 ²⁴	Bibliography BOT - BIBOT	Localização de estudos em Bases de Dados	Não Atende
Popoff, Besada, Jansen, et al. 2020 ²⁵	SVM; NB; CART	Seleção 1ª triagem	Atende
Burns, Etherington, Cheng-Boivin, Boet, 2021 ²⁶	DistillerAI	Seleção 1ª triagem	Atende
Chai, Lines, Gucciardi, Ng, 2021 ²⁷	Research Screener	Seleção 1ª triagem	Atende
Pham, Jovanovic, Bagheri, Antony, et al. 2021 ²⁸	Mineração de texto; LDA; SVD; PLN	Seleção 1ª triagem	Atende
Qin, Liu, Wang, et al. 2021 ²⁹	PNL; BERT; LightGBM	Seleção 1ª triagem	Atende
Borissov, Haas, Minder, et al. 2022 ³⁰	Deduklick	Retirada de duplicados	Atende
Facchinetti, Benetti, Giuffrida, Nocera, 2022 ³¹	Slr-kit	Seleção 1ª triagem	Não Atende
Muller, Ames, Jardim, Rose, 2022 ³²	Lingo3G	Seleção 1ª triagem	Atende
Reis, Oliveira, Fritsch, et al. 2023 ³³	Rayyan; Abstrackr; Colandr	Seleção 1ª triagem	Atende
Kebede, Cornet, Fortner, 2023 ³⁴	NB; SVM; SVD	Seleção 1ª triagem	Não Atende
Li, Kabouji, Bouhadoun, et al. 2023 ³⁵	Rayyan; Abstrackr; SWIFT-Review	Seleção 1ª triagem	Não Atende
Natukunda, Muchene, 2023 ³⁶	Latent Dirichlet Allocation	Seleção 1ª triagem	Não Atende
Oude Wolcherink, Pouwels, van Dijk, et al. 2023 ³⁷	ASReview	Seleção 1ª triagem	Não Atende
Qureshi, Shaughnessy, Gill, et al. 2023 ³⁸	ChatGPT	Pergunta e Estratégia de busca	Não atende
van Dijk Brusse-Keizer, Bucsán, et al. 2023 ³⁹	ASReview	Seleção 1ª triagem	Atende

FONTE: Os autores, 2025.

A análise dos 29 artigos revelou que 51,7% (15 estudos) atenderam aos manuais JBI e Cochrane em pelo menos uma etapa da RS, enquanto os 48,3% (14 estudos) não atenderam. A adequação das aplicações variou consideravelmente conforme a etapa da RS em que foram utilizadas, destacando diferenças na implementação e alinhamento com os manuais.

A maioria dos estudos analisados aplicou ferramentas de IA na fase de Seleção de Estudos (1ª Triagem), representando 26 dos 29 estudos (89,7%). Desses 26, metade (13 aplicações) demonstrou alinhamento com os manuais, enquanto a outra metade falhou em atender aos critérios metodológicos estabelecidos. Entre os estudos que se destacaram positivamente estão: Burns et al.²⁶, com aplicação do DistillerAI;

e Rathbone et al.¹⁵, com o Abstrackr. Ambos estudos atenderam os critérios na triagem inicial. Em contrapartida, os estudos de Bucheli Guerrero¹⁹, com o SASR; e Oude Wolcherink et al.³⁷, com o ASReview, não atenderam aos manuais devido a deficiências no cumprimento de critérios metodológicos, como a falta de justificativas rastreáveis para exclusões de estudos.

Na etapa de Retirada de Duplicados,

apenas um estudo foi avaliado: Borissov et al.³⁰, com o Deduklick. Este demonstrou alinhamento aos manuais. Para a fase de Formulação de Pergunta de Pesquisa e Estratégia de Busca, o estudo de Qureshi et al.³⁸, com o ChatGPT, foi avaliado, mas não atendeu às recomendações dos manuais.

A análise cronológica revelou uma evolução significativa nos recursos de IA ao longo dos anos. No período de 2012 a 2016, as tecnologias iniciais apontadas nesses estudos, como o SVM (Support Vector Machine), um algoritmo de aprendizado supervisionado projetado para classificar dados, como as redes neurais – estruturas computacionais inspiradas no funcionamento do cérebro humano e capazes de modelar padrões complexos por meio de camadas de processamento interconectadas – não apresentaram alinhamento com os critérios metodológicos estabelecidos. Entre 2018 e 2021, observou-se um aumento na proporção de softwares alinhados aos manuais. Destacam-se estudos como os de Przybyła et al.¹⁷, com o Robot Analyst, e Burns et al.²⁶, com o DistillerAI. Ambos atenderam aos critérios metodológicos nas etapas, como a 1ª triagem e a 2ª triagem. No período mais recente, entre 2022 e 2023, o estudo de Van Dijk et al.³⁹ destacou-se pela aplicação do ASReview, que demonstrou atender os manuais. No entanto, quando o mesmo software foi utilizado por Oude Wolcherink et al.³⁷, a ausência de justificativas rastreáveis para exclusões de estudos comprometeu sua transparência e, conseqüentemente, sua conformidade com os manuais.

De forma geral, os estudos que mais se destacaram foram aqueles que utilizaram ferramentas de IA em etapas operacionais, como a triagem inicial e a remoção de duplicados, apresentando alinhamento com os manuais. Por outro lado, etapas relacionadas ao planejamento e à formulação de estratégias de busca mostraram-se problemáticas, refletindo limitações tanto nas ferramentas empregadas quanto na forma como

foram implementadas pelos pesquisadores.

DISCUSSÃO

Os resultados apresentados neste trabalho evidenciam que, embora o uso de aplicações baseadas em IA em RS esteja avançando, apenas metade dos estudos avaliados atende às diretrizes metodológicas JBI e Cochrane. Esse cenário revela um contraste entre as etapas operacionais, nas quais a aplicação de IA tem se mostrado eficaz, e as etapas críticas de planejamento e análise, que ainda dependem de supervisão humana.

A fase de seleção, especificamente a primeira triagem, foi contemplada em 23 dos 29 estudos avaliados, Rathbone et al.¹⁵, com a aplicação do Software Abstrackr, Gartlehner et al.²⁰ com o DistillerAI e Li et al.³⁵ com o SWIFT-Review demonstraram eficácia ao reduzir significativamente a carga de trabalho manual.

Por outro lado, o estudo de Przybyła et al.¹⁷, com o software Robot Analyst, atendeu aos manuais, porém, no cenário automatizado, excluiu até 70% dos registros relevantes, o que pode comprometer a confiabilidade dos resultados das RS. Já o estudo de Gartlehner et al.²⁰, com o software DistillerAI, embora tenha sido eficiente em reduzir a carga de trabalho (99%), apresentou elevada perda de registros (97%) no cenário automatizado. Esses dados reforçam que a supervisão humana continua essencial para garantir a qualidade e a integridade do processo de seleção.

Programas e plataformas que utilizam algoritmos avançados, como o AS-Review, citado no estudo de Van Dijk et al.³⁹, mostraram eficiência na seleção da 1ª triagem ao combinar modelos como Naïve Bayes e TF-IDF. Já no estudo de Oude Wolcherink et al.³⁷ não foram apresentados critérios claros para as exclusões de trabalhos, o que dificulta a replicação e a validação dos resultados, comprometendo sua aplicação.

A deduplicação foi uma das etapas

com maior sucesso no uso de IA. Borissov et al.³⁰ apresentou o Deduklick e Reis et al.³³ apresentou o Rayyan, ambos garantiram rastreabilidade e padronização, identificando e eliminando duplicatas de forma automática e eficaz. O Deduklick³⁰ destacou-se por utilizar medidas de similaridade e gerar relatórios no formato PRISMA.

Apesar dos avanços, as aplicações de IA avaliadas para busca automatizada em bases de dados, como o BiBot estudados por Orgeolet et al.²⁴, não atenderam aos manuais. Limitações como a ausência de descritores autorizados da área da saúde, como MeSH, DeCS e de operadores booleanos, comprometeram a sensibilidade e a especificidade das estratégias de busca.

As etapas como formulação de estratégia de busca, avaliação de risco de vies e síntese de resultados não foram abordadas nos artigos analisados neste trabalho. Isso reflete uma brecha significativa na aplicação da IA em RS, uma vez que essas etapas exigem julgamento crítico e contextual, características que as tecnologias atuais ainda não replicam adequadamente.

CONCLUSÃO

Este artigo buscou identificar como as RS que utilizam ferramentas baseadas em IA na área da saúde se alinham aos manuais JBI e Cochrane, os quais estabelecem diretrizes metodológicas para garantir a transparência, a reprodutibilidade e a qualidade das revisões. Os resultados indicaram que, apesar dos avanços nas aplicações de IA, especialmente em etapas operacionais, como a triagem inicial e deduplicação de registros, apenas 15 dos 29 estudos atenderam os manuais.

As aplicações que atenderam as diretrizes metodológicas dos manuais destacaram-se por sua rastreabilidade, padronização e transparência. Exemplo disso foi o trabalho de Borissov et al.³⁰, com aplicação da ferramenta Deduklick, e de Pham et al.²⁸, com a mineração

de texto, os quais exemplificam como a IA pode ser aplicada com sucesso em etapas operacionais das RS. Por outro lado, estudos como Oude Wolcherink³⁷, com o ASReview, e Orgeolet²⁴, com BIBOT, apresentaram falhas por não documentarem os critérios adotados nas decisões tomadas.

Esses achados mostram que as limitações observadas não estão intrinsecamente ligadas às aplicações de IA, mas à maneira como elas são empregadas pelos pesquisadores. A falta de adesão às diretrizes dos manuais JBI e Cochrane compromete a confiabilidade das RS, mesmo quando ferramentas tecnológicas avançadas são utilizadas. A ausência de estratégias padronizadas, a utilização limitada das funcionalidades das ferramentas e a dependência de decisões pouco transparentes comprometem a validade e a reprodutibilidade dos resultados. Para que as RS apoiadas por

IA sejam consistentes e úteis na prática baseada em evidências, é imprescindível que as ferramentas sejam compatíveis com as diretrizes estabelecidas pelos manuais.

Portanto, o objetivo deste estudo foi alcançado. Os resultados permitiram identificar que as aplicações baseadas em IA atendem os manuais JBI e Cochrane, mas esse alinhamento ainda é limitado. Isso ocorre tanto pelas limitações das ferramentas quanto pelo uso inadequado delas pelos pesquisadores.

Como contribuição, esta revisão destaca a necessidade de conscientizar os pesquisadores sobre as diretrizes metodológicas dos manuais, bem como de um esforço conjunto entre desenvolvedores de softwares e a comunidade científica. Isso traria como objetivo a promoção de soluções robustas e integradas, melhorias nas funcionalidades das aplicações de IA e o treinamento

dos usuários para sua aplicação metodologicamente correta. Essas medidas são consideradas essenciais para que essas ferramentas cumpram seu papel no fortalecimento das RS, garantindo maior eficiência sem comprometer a qualidade científica.

Este estudo foi realizado com base em RS publicadas, o que restringe os achados às aplicações baseadas em IA descritas nessas fontes. A heterogeneidade nos objetivos, metodologias e dados reportados pelos estudos avaliados dificulta a generalização dos resultados, uma vez que nem todas as aplicações foram testadas em condições uniformes, o que pode ter limitado a compreensão detalhada das tecnologias. Apesar dessas limitações, este trabalho fornece uma base importante para discutir os avanços, desafios e potencialidades do uso de IA em RS na área da saúde.

REFERÊNCIAS

1. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JMM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;29(372). Disponível em: <https://doi.org/10.1136/bmj.n71>
2. Clarke M, Chalmers I. Reflections on the history of systematic reviews. *BMJ Evid Based Med*. 2018;23(4):121-122. Disponível em: <https://doi.org/10.1136/bmjebm-2018-110968>
3. Ahn HS, Kim HJ. An introduction to systematic review. *J Korean Med Assoc*. 2014;57(1):49-59. Disponível em: https://www.researchgate.net/publication/272396557_An_introduction_to_systematic_review
4. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, Editors. *JBI Manual for Evidence Synthesis*. JBI; 2024. Disponível em: <https://synthesismanual.jbi.global>
5. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., Editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.5, 2024. Disponível em: <https://training.cochrane.org/handbook/current>
6. Gusenbauer M, Haddaway NR. What every researcher should know about searching – Clarifying the fundamentals of literature searching. *Res Synth Methods*. 2020;11(3):136-147. Disponível em: <https://doi.org/10.1002/jrsm.1427>
7. Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Quarterly*. 2016;94(3):485-514. Disponível em: <https://doi.org/10.1111/1468-0009>
8. Blaizot A, Veettil SK, Saidoung P, Moreno-Garcia CF, Wiratunga N, Aceves-Martins M, Lai NM, Chaiyakunapruk N. Using artificial intelligence methods for systematic review in health sciences: a systematic review. *Res Synth Methods*. 2022;13(3):353-362. Disponível em: <https://doi.org/10.1002/jrsm.1553>
9. Tóth B, Berek I, Gulácsi I, et al. automation

- of systematic reviews of biomedical literature: a scoping review of studies indexed in Pubmed. *Syst Rev.* 2024;13:174. Disponível em: <https://doi.org/10.1186/s13643-024-02592-3>.
10. Gil AC. Como elaborar projetos de pesquisa. 6. ed. São Paulo: Atlas; 2017.
11. Bekhuis T, Demner-Fushman D. Screening non-randomized studies for medical systematic reviews: a comparative study of classifiers. *Artif Intell Med.* 2012;55(3):197-207. Disponível em: <https://doi.org/10.1016/j.artmed.2012.05.002>
12. Jonnalagadda S, Petitti D. A new iterative method to reduce workload in systematic review process. *Int J Comput Biol Drug Des.* 2013;6(1-2):5-17. Disponível em; <https://doi.org/10.1504/ijcbdd.2013.052198>
13. Kim S, Choi J. Improving the performance of text categorization models used for the selection of high quality articles. *Healthc Inform Res.* 2012;18(1):18-28. Disponível em: <https://doi.org/10.4258/hir.2012.18.1.18>
14. Blake C, Lucic A. Automatic endpoint detection to support the systematic review process. *J Biomed Inform.* 2015 Aug;56:42-56. Disponível em: <https://doi.org/10.1016/j.jbi.2015.05.004>
15. Rathbone J, Hoffmann T, Glasziou P. Faster title and abstract screening? Evaluating Abstrackr, a semi-automated online screening program for systematic reviewers. *Syst Rev.* 2015;4:80. Disponível em: <https://doi.org/10.1186/s13643-015-0067-6>
16. Hashimoto K, Kontonatsios G, Miwa M, Ananiadou S. Topic detection using paragraph vectors to support active learning in systematic reviews. *J Biomed Inform.* 2016;62:59-65. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC4981645/>
17. Przybyła P, Brockmeier AJ, Kontonatsios G, Le Pogam MA, McNaught J, von Elm E, Nolan K, Ananiadou S. Prioritising references for systematic reviews with RobotAnalyst: a user study. *Res Synth Methods.* 2018;9(3):470-488. Disponível em: <https://doi.org/10.1002/jrsm.1311>
18. Tsafnat G, Glasziou P, Karystianis G, Coiera E. Automated screening of research studies for systematic reviews using study characteristics. *Syst Rev.* 2018 Apr 25;7(1):64. Disponível em: <https://doi.org/10.1186/s13643-018-0724-7>
19. Bucheli Guerrero VA. Desarrollo del estado del arte en investigación: una herramienta basada en inteligencia artificial. *Rev Politécnica.* 2019;15(30):70-81. Disponível em; <http://dx.doi.org/10.33571/rpolitec.v15n30a7>
20. Gartlehner G, Wagner G, Lux L, Affengruber L, Dobrescu A, Kaminski-Hartenthaler A, Viswanathan M. Assessing the accuracy of machine-assisted abstract screening with DistillerAI: a user study. *Syst Rev.* 2019;8(1):277. Disponível em: <https://doi.org/10.1186/s13643-019-1221-3>
21. Gates A, Guitard S, Pillay J, Elliott SA, Dyson MP, Newton AS, Hartling L. Performance and usability of machine learning for screening in systematic reviews: a comparative evaluation of three tools. *Syst Rev.* 2019;8(1):278. Disponível em; <https://doi.org/10.1186/s13643-019-1222-2>
22. Gates A, Gates M, Sebastianski M, Guitard S, Elliott SA, Hartling L. The semi-automation of title and abstract screening: a retrospective exploration of ways to leverage Abstrackr's relevance predictions in systematic and rapid reviews. *BMC Med Res Methodol.* 2020;20(1):139. Disponível em: <https://doi.org/10.1186/s12874-020-01031-w>
23. Howard BE, Phillips J, Tandon A, Maharana A, Elmore R, Mav D, Sedykh A, Thayer K, et al. SWIFT-Active Screener: accelerated document screening through active learning and integrated recall estimation. *Environ Int.* 2020;138:105623. Disponível em: <https://doi.org/10.1016/j.envint.2020.105623>
24. Orgeolet L, Foulquier N, Misery L, Redou P, Pers JO, Devauchelle-Pensec V, Saraux A. Can artificial intelligence replace manual search for systematic literature? Review on cutaneous manifestations in primary Sjögren's syndrome. *Rheumatology (Oxford).* 2020;59(4):811-819. Disponível em; <https://doi.org/10.1093/rheumatology/kez370>
25. Popoff E, Besada M, Jansen JP, Cope S, Kanters S. Aligning text mining and machine learning algorithms with best practices for study selection in systematic literature reviews. *Syst Rev.* 2020 Dec 13;9(1):293. Disponível em: <https://doi.org/10.1186/s13643-020-01031-w>

org/10.1186/s13643-020-01520-5

26. Burns JK, Etherington C, Cheng-Boivin O, Boet S. Using an artificial intelligence tool can be as accurate as human assessors in level one screening for a systematic review. *Health Info Libr J*. 2024;41(2):136-148. Disponível em: <https://doi.org/10.1111/hir.12413> Epub 2021. PMID: 34792285.

27. Chai KEK, Lines RLJ, Gucciardi DF, Ng L. Research Screener: a machine learning tool to semi-automate abstract screening for systematic reviews. *Syst Rev*. 2021;10(1):93. Disponível em: <https://doi.org/10.1186/s13643-021-01635-3>

28. Pham B, Jovanovic J, Bagheri E, Antony J, Ashoor H, Nguyen TT, Rios P, Robson R, Thomas SM, Watt J, Straus SE, Tricco AC. Text mining to support abstract screening for knowledge syntheses: a semi-automated workflow. *Syst Rev*. 2021;10(1):156. Disponível em: <https://doi.org/10.1186/s13643-021-01700-x>.

29. Qin X, Liu J, Wang Y, Liu Y, Deng K, Ma Y, Zou K, Li L, Sun X. Natural language processing was effective in assisting rapid title and abstract screening when updating systematic reviews. *J Clin Epidemiol*. 2021;133:121-129. Disponível em: <https://doi.org/10.1016/j.jclinepi.2021.01.010>

30. Borissov N, Haas Q, Minder B, Kopp-Heim D, von Gernler M, Janka H, Teodoro D, Amini P. Reducing systematic review burden using Deduklick: a novel, automated, reliable, and explainable deduplication algorithm to foster medical research. *Syst Rev*. 2022;11(1):172. Disponível em: <https://doi.org/10.1186/s13643-022-02045-9>.

31. Facchinetti T, Benetti G, Giuffrida D, Nocera A. Slr-kit: a semi-supervised machine learning framework for systematic literature reviews. *Knowledge-Based Systems*. 2022;251:109266. Disponível em: <https://doi.org/10.1016/j.knsys.2022.109266>.

32. Muller AE, Ames HMR, Jardim PSJ, Rose CJ. Machine learning in systematic reviews: Comparing automated text clustering with Lingo3G and human researcher categorization in a rapid review. *Res Synth Methods*. 2022 Mar;13(2):229-241. Disponível em: <https://doi.org/10.1002/jrsm.1541>.

33. Reis AHS, de Oliveira ALM, Fritsch C, Zouch J, Ferreira P, Polese JC. Usefulness of machine learning softwares to screen titles of systematic reviews: a methodological study. *Syst Rev*. 2023;12(1):68. Disponível em: <https://doi.org/10.1186/s13643-023-02231-3>.

34. Kebede MM, Le Cornet C, Fortner RT. In-depth evaluation of machine learning methods for semi-automating article screening in a systematic review of mechanistic literature. *Res Synth Methods*. 2023 Mar;14(2):156-172. Disponível em: <https://doi.org/10.1002/jrsm.1589>.

35. Li J, Kabouji J, Bouhadoun S, Tanveer S, Filion KB, Gore G, Josephson CB, Kwon CS, et al. Sensitivity and specificity of alternative screening methods for systematic reviews using text mining tools. *J Clin Epidemiol*. 2023;162:72-80. Disponível em: <https://doi.org/10.1016/j.jclinepi.2023.07.010>.

36. Natukunda A, Muchene LK. Unsupervised title and abstract screening for systematic review: a retrospective case-study using topic modelling methodology. *Syst Rev*. 2023 Jan 3;12(1):1. Disponível em: <https://doi.org/10.1186/s13643-022-02163-4>.

37. Oude Wolcherink MJ, Pouwels XGLV, van Dijk SHB, Doggen CJM, Koffijberg H. Can artificial intelligence separate the wheat from the chaff in systematic reviews of health economic articles? *Expert Rev Pharmacoecon Outcomes Res*. 2023;23(9):1049-1056. Disponível em: <https://doi.org/10.1080/14737167.2023.2234639>.

38. Qureshi R, Shaughnessy D, Gill KAR, Robinson KA, Li T, Agai E. Are ChatGPT and large language models "the answer" to bringing us closer to systematic review automation? *Syst Rev*. 2023 Apr 29;12(1):72. Disponível em: <https://doi.org/10.1186/s13643-023-02243-z>.

39. Van Dijk SHB, Brusse-Keizer MGJ, Bucsan CC, Van der Palen J, Doggen CJM, Lenferink A. Artificial intelligence in systematic reviews: promising when appropriately used. *BMJ Open*. 2023;13:e072254. Disponível em: <https://doi.org/10.1136/bmjopen-2023-072254>.