OBJN
Online Brazilian Journal of Nursing
ISSN: 1676 - 4285 | www.objnursing.uff.br

Submissão: 28-Mai-2026
Aprovado: 26-Ago-2026

SELEÇÃO DE ESTUDOS ASSISTIDA PELA INTELIGÊNCIA ARTIFICIAL CLAUDE AI: UM GUIA METODOLÓGICO

STUDY SELECTION ASSISTED BY CLAUDE AI ARTIFICIAL INTELLIGENCE: A METHODOLOGICAL GUIDE

SELECCIÓN DE ESTUDIOS ASISTIDA POR LA INTELIGENCIA ARTIFICIAL CLAUDE AI: UNA GUÍA METODOLÓGICA

Rômulo Severo Sampaio[1] ORCID: 0000-0002-0657-2646
Yasmim Ribeiro Fracaroli[2] ORCID: 0009-0004-2400-2323
Anna Paula Mendes Marques de Lima Franco[2] ORCID: 0009-0009-8847-7914
Marianne Alves Rodrigues[2] ORCID: 0009-0001-6504-9089
Francini Castilha do Nascimento[2] ORCID: 0000-0001-7860-0493
Cecília Santos de Lima[2] ORCID: 0009-0004-3706-2368
Isabelle Cristinne Pinto Sampaio Costa[2] ORCID: 0000-0002-2611-8643

[1] Universidade Professor Edson Antônio Velano. Alfenas, Minas Gerais, Brasil.
[2] Universidade Federal de Alfenas (UNIFAL-MG). Alfenas, Minas Gerais, Brasil.

Como citar: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, et al. Study selection assisted by Claude AI artificial intelligence: a methodological guide. Online Braz J Nurs. 2026;25(Suppl1):e20267039. https://doi.org/10.17665/1676-4285.20267039

RESUMO

Objetivo: Apresentar um guia metodológico para a seleção de estudos em revisões de literatura com o auxílio do Claude AI. **Método:** Estudo de reflexão teórico-metodológica, fundamentado em revisão narrativa da literatura e na experiência empírica dos autores na condução de sínteses de evidências. **Resultados:** Propõe-se um fluxo metodológico estruturado em cinco etapas: (1) exportação e deduplicação inicial no EndNote; (2) deduplicação adicional e fragmentação do corpus em blocos anuais no Rayyan; (3) criação de um projeto no Claude AI, incluindo os critérios de elegibilidade; (4) triagem de títulos e resumos assistida por inteligência artificial (IA), com geração automatizada de planilha contendo a decisão e a justificativa para cada registro; e (5) retorno ao Rayyan para revisão humana independente e posterior integração ao fluxograma PRISMA. Esse processo configura um modelo teórico-operacional que fornece recomendações práticas para a definição dos critérios de elegibilidade, elaboração estruturada de prompts e registro transparente das decisões. **Conclusão:** O guia estabelece uma estrutura inovadora para a incorporação da IA na etapa de seleção de estudos em revisões de literatura. Essa abordagem acelera as fases iniciais da síntese de evidências e orienta o uso responsável da tecnologia. Como limitação, aponta-se a ausência de validação empírica multicêntrica do fluxo proposto, o que demanda sua aplicação e avaliação em pesquisas futuras.

Descritores: Inteligência Artificial; Inteligência Artificial Generativa; Prática Clínica Baseada em Evidências; Metodologia como Assunto; Pesquisa; Revisão Sistemática.

ABSTRACT

Objective: To present a methodological guide for study selection in literature reviews assisted by Claude AI. **Method:** A theoretical-methodological reflection study grounded in a narrative literature review and the authors' empirical experience in conducting evidence syntheses. **Results:** A five-stage structured methodological workflow is proposed: (1) export and initial deduplication in EndNote; (2) additional deduplication and corpus segmentation into annual blocks in Rayyan; (3) creation of a project in Claude AI incorporating the eligibility criteria; (4) AI-assisted title and abstract screening with automated generation of a spreadsheet containing the decision and justification for each record; and (5) return to Rayyan for independent human review and subsequent integration into the PRISMA flow diagram. This process establishes a theoretical-operational model that provides practical recommendations for defining eligibility criteria, developing structured prompts, and transparently documenting decisions. **Conclusion:** The guide establishes an innovative framework for incorporating AI into the study selection phase of literature reviews. This approach accelerates the initial stages of evidence synthesis and promotes the responsible use of technology. A noted limitation is the lack of multicenter empirical validation of the proposed workflow, which warrants its application and evaluation in future research.

Descriptors: Artificial Intelligence; Generative Artificial Intelligence; Evidence-Based Practice; Methodology as a Subject; Research; Systematic Review.

RESUMEN

Objetivo: Presentar una guía metodológica para la selección de estudios en revisiones de literatura con el auxilio de Claude AI. **Método:** Estudio de reflexión teórico-metodológica, fundamentado en una revisión narrativa de la literatura y en la experiencia empírica de los autores en la conducción de síntesis de evidencias. **Resultados:** Se propone un flujo metodológico estructurado en cinco etapas: (1) exportación y deduplicación inicial en EndNote; (2) deduplicación adicional y fragmentación del corpus en bloques anuales en Rayyan; (3) creación de un proyecto en Claude AI, incluyendo los criterios de elegibilidad; (4) tamizaje de títulos y resúmenes asistido por inteligencia artificial (IA), con generación automatizada de una hoja de cálculo que contiene la decisión y la justificación para cada registro; y (5) retorno a Rayyan para la revisión humana independiente y su posterior integración al diagrama de flujo PRISMA. Este proceso configura un modelo teórico-operacional que proporciona recomendaciones prácticas para la definición de los criterios de elegibilidad, la elaboración estructurada de prompts y el registro transparente de las decisiones. **Conclusión:** La guía establece una estructura innovadora para la incorporación de la IA en la etapa de selección de estudios en revisiones de literatura. Este enfoque acelera las fases iniciales de la síntesis de evidencias y orienta el uso responsable de la tecnología. Como limitación, se señala la ausencia de validación empírica multicéntrica del flujo propuesto, lo que demanda su aplicación y evaluación en futuras investigaciones.

Descriptores: Inteligencia Artificial; Inteligencia Artificial Generativa; Práctica Clínica Basada en la Evidencia; Metodología como un Tema; Investigación; Revisión Sistemática.

Editores:
Rosimere Ferreira Santana (ORCID: 0000-0002-4593-3715)
Geilsa Soraia Cavalcanti Valente (ORCID: 0000-0003-4488-4912)
Emerson Willian Santos de Almeida (ORCID: 0000-0002-6846-021X)

Editora:
Escola de Enfermagem Aurora de Afonso Costa – UFF
Rua Dr. Celestino, 74 – Centro, CEP: 24020-091 – Niterói, RJ, Brasil
E-mail da revista: objn.cme@id.uff.br

Autora correspondente:
Yasmim Ribeiro Fracaroli
E-mail: yasmim.fracaroli@sou.unifal-mg.edu.br

O que já se sabe:

- Os métodos tradicionais de triagem em revisões exigem elevado investimento de tempo e apresentam limitações operacionais no processo de síntese de evidências.
- A IA generativa possui potencial para acelerar a seleção de estudos, por título e resumo, em pesquisas na área da saúde.
- Ainda existem lacunas metodológicas referentes à padronização, à transparência e à validação do uso de IA em revisões sistemáticas.

O que este artigo acrescenta:

- O uso do Claude AI pode reduzir o tempo despendido na triagem inicial de estudos para revisões de literatura em saúde.
- O fluxo metodológico proposto amplia a transparência e a padronização das decisões durante a seleção de estudos assistida por IA.
- A integração entre Claude AI, Rayyan, EndNote e PRISMA favorece maior rastreabilidade e organização no processo de síntese de evidências.

INTRODUÇÃO

Os estudos de revisão são instrumentos centrais na Prática Baseada em Evidências (PBE), pois permitem reunir e analisar criticamente o conhecimento disponível sobre determinada temática. Desse modo, oferecem suporte qualificado para a tomada de decisões clínicas e para a formulação de políticas de saúde. Tais estudos sintetizam o saber de forma confiável, favorecem a detecção de lacunas no campo investigado e orientam a definição de novas questões de pesquisa[1-3].

A triagem de artigos por meio da leitura de títulos e resumos constitui uma das etapas mais exaustivas e demoradas da revisão de literatura, com impacto direto na validade e na reprodutibilidade dos resultados. Estimativas apontam que uma revisão sistemática convencional consome, em média, de seis a 18 meses de trabalho. A fase de seleção de estudos responde por uma parcela substancial desse período, tornando-se particularmente vulnerável à fadiga humana e à deriva interpretativa entre os revisores[4,5].

Gerenciadores de referências, como o EndNote, e plataformas dedicadas à triagem colaborativa, como o Rayyan, consolidaram-se como recursos indispensáveis no fluxo de revisões. Embora essas tecnologias automatizem a deduplicação e a organização do *corpus* de manuscritos identificados nas buscas, a decisão de elegibilidade permanece centrada na leitura, na interpretação e na cognição humana, com a avaliação de cada registro individualmente[6,7].

A Inteligência Artificial (IA) tem transformado os estudos de revisão ao automatizar etapas como busca, triagem e síntese de dados, conferindo maior celeridade e precisão ao processo metodológico[8]. Mais recentemente, destacam-se os *Large Language Models* (LLMs), sistemas capazes de compreender, gerar e estruturar textos complexos, o que amplia significativamente as possibilidades de suporte tecnológico à seleção de estudos nessas pesquisas[9].

Modelos como o ChatGPT (OpenAI), o Gemini (Google) e o Claude AI (Anthropic) demonstraram, em avaliações preliminares, desempenho promissor na triagem de títulos e resumos. Em domínios específicos, essas ferramentas alcançaram sensibilidade e especificidade comparáveis às de revisores humanos[10,11]. Contudo, a validação humana permanece imperativa para assegurar a qualidade metodológica, mitigar vieses e garantir interpretações contextualizadas[8,12].

Entre as tecnologias disponíveis, o Claude AI apresenta particularidades que o tornam singularmente atrativo para revisões na área da saúde. A ferramenta possui ampla janela de contexto (até 200 mil *tokens*), suporta a criação de projetos com arquivos persistentes anexados e processa extensos lotes de registros bibliográficos em uma única interação. Além disso, gera justificativas auditáveis, em linguagem natural, para cada decisão de inclusão ou exclusão de artigos[13]. Tais recursos viabilizam a manutenção de critérios de elegibilidade estáveis ao longo da triagem e conferem rastreabilidade ao processo. Esses atributos são fundamentais para a transparência preconizada pelas diretrizes *Preferred Reporting Items for Systematic reviews and Meta-Analyses* (PRISMA) e pelos manuais *Cochrane Handbook* e JBI[14-16].

A despeito dos avanços recentes, persiste uma lacuna na literatura quanto à ausência de orientações metodológicas sistematizadas para integrar — de forma estruturada, crítica e ética — o uso de LLMs à etapa de seleção de estudos. A produção científica atual tende a concentrar-se na comparação de desempenho entre a IA e revisores humanos ou em discussões teóricas sobre os riscos da automação. Consequentemente, carece-se de um guia prático que articule as etapas intermediárias — como gerenciamento de referências, deduplicação, fragmentação do *corpus* e configuração da ferramenta — à decisão final de elegibilidade dos estudos[16-18].

Diante desse cenário, este estudo objetiva apresentar um guia metodológico para a seleção de estudos em revisões de literatura com o auxílio de IA generativa — especificamente do modelo Claude AI —, com foco na otimização da síntese de evidências na área da saúde.

MÉTODO

Trata-se da proposição de um guia metodológico fundamentado em reflexão teórico-metodológica. O processo foi conduzido por meio de uma revisão narrativa da literatura sobre o uso de IA generativa na síntese de evidências, aliada à experiência dos autores na condução de diferentes tipos de revisões (sistemáticas, integrativas e de escopo). O desenvolvimento do guia orientou-se pelas diretrizes internacionais PRISMA[14], *Cochrane Handbook*[15] e *JBI Manual for Evidence Synthesis*[16], além de adotar os preceitos reconhecidos para a elaboração de manuais orientadores em saúde[19].

A reflexão ocorreu de forma estruturada em cinco etapas: (1) levantamento e análise crítica de publicações metodológicas e de estudos empíricos sobre a aplicação de LLMs na triagem de estudos[8,12,17,18]; (2) mapeamento das funcionalidades dessas ferramentas em relação aos requisitos da diretriz PRISMA para a triagem de títulos e resumos; (3) análise comparativa entre as recomendações

metodológicas de organizações de referência internacional (como Cochrane e JBI) e as potencialidades e limitações operacionais do Claude AI; (4) sistematização da experiência prática dos autores na aplicação iterativa do fluxo proposto, visando identificar oportunidades, riscos metodológicos e requisitos para o uso responsável da ferramenta; e (5) consolidação das etapas identificadas em um fluxo metodológico, culminando na proposição do guia e do fluxograma operacional expostos na seção de Resultados.

A busca na literatura foi realizada nas fontes de informação *Medical Literature Analysis and Retrieval System Online* (PubMed/MEDLINE), Scopus, Web of Science Core Collection (WoS) e Google Scholar. Estabeleceu-se o recorte temporal de 2020 a 2025, justificado pela recência do lançamento comercial dos LLMs. Utilizaram-se os descritores "*artificial intelligence*", "*large language model*", "*systematic review*", "*screening*", "*title and abstract*" e "Claude AI", bem como seus correspondentes em português, cruzados mediante os operadores booleanos *AND* e *OR*. Foram incluídos estudos publicados em inglês, português ou espanhol, sem restrição de desenho metodológico, que abordassem o uso de IA na síntese de evidências. Excluíram-se os artigos sem descrição metodológica clara ou que divergiam do escopo temático.

Por tratar-se de uma revisão narrativa, de caráter reflexivo e não sistemático, prescindiu-se do registro prévio de protocolo em plataformas como o *International Prospective Register of Systematic Reviews* (PROSPERO) ou o *Open Science Framework* (OSF), em conformidade com a natureza do desenho do estudo[17].

A triagem da literatura ocorreu em duas fases: leitura de títulos e resumos, seguida da avaliação na íntegra dos textos potencialmente relevantes. A análise dos dados foi conduzida de forma interpretativa, com foco na identificação de estratégias, limitações e recomendações para o uso da IA na seleção de artigos. Priorizaram-se as experiências aplicáveis ao contexto da saúde, com ênfase em propostas metodológicas e estudos de validação empírica.

A experiência empírica dos autores foi incorporada de forma sistematizada a partir da aplicação prática e da reflexão crítica sobre o uso do Claude AI durante a condução de uma revisão sistemática focada em protocolos de dessensibilização a taxanos em mulheres com câncer de mama. Tal pesquisa encontra-se em andamento e está devidamente registrada no PROSPERO (CRD420261325732). A escolha desse estudo como cenário de aplicação decorreu de sua sinergia com a linha de pesquisa do grupo, viabilizando o teste e o refinamento do guia em um contexto real de produção científica. Durante a testagem, processaram-se blocos anuais de registros bibliográficos — previamente deduplicados no EndNote e no Rayyan —, inserindo-se os critérios de elegibilidade como arquivos de contexto persistentes em um projeto dedicado no Claude AI. A equipe de pesquisa analisou criticamente as decisões e as justificativas geradas pelo modelo, o que propiciou o aprimoramento iterativo do método aqui documentado.

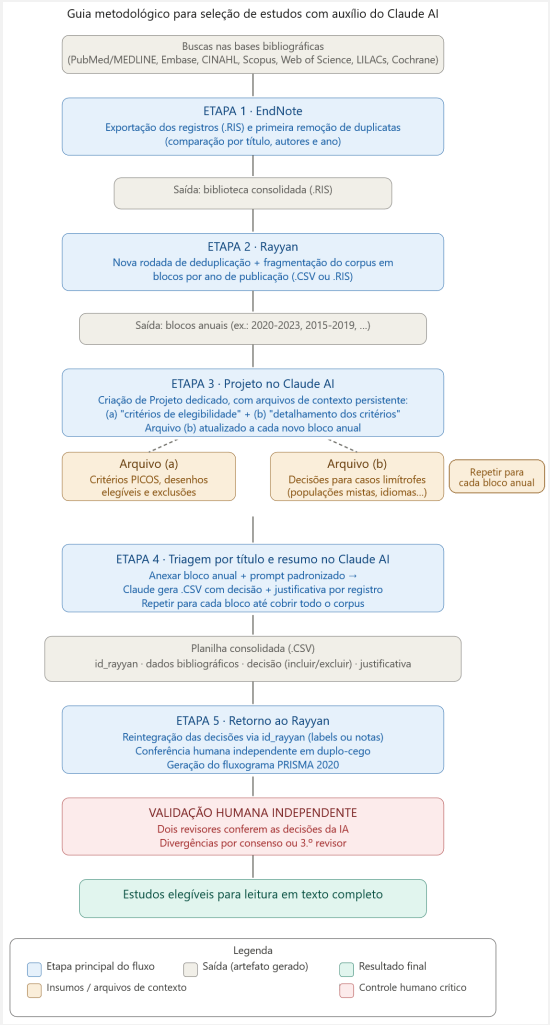
No tocante à segurança e à privacidade dos dados, a aplicação do fluxo ocorreu mediante o uso da versão Claude Pro (interface claude.ai). Desativou-se a configuração de privacidade "*Help Improve Claude*" para impedir que os comandos (*prompts*) e os arquivos anexados fossem utilizados no treinamento de modelos futuros da empresa Anthropic[20]. O fornecimento de informações à ferramenta restringiu-se aos metadados bibliográficos dos registros (identificador, título, autores, ano, periódico, DOI, resumo e palavras-chave). Não foram submetidos textos na íntegra, dados inéditos ou quaisquer outros materiais sensíveis.

A proposição metodológica resultante estruturou-se em cinco etapas sequenciais, descritas na seção de Resultados e ilustradas em um fluxograma operacional. Por tratar-se de estudo de natureza teórico-metodológica, sem envolvimento direto de seres humanos, dispensou-se a apreciação por Comitê de Ética em Pesquisa (CEP), respeitando-se integralmente as exigências da Lei nº 14.874/2024 e das Resoluções nº 466/2012 e nº 510/2016 do Conselho Nacional de Saúde (CNS).

RESULTADOS

O guia metodológico proposto estrutura-se em cinco etapas sequenciais, articulando ferramentas tradicionais de gerenciamento de referências e de triagem colaborativa (EndNote e Rayyan) ao uso do Claude AI como recurso auxiliar na decisão de elegibilidade por título e resumo. A Figura 1 apresenta a representação esquemática do fluxo.

Figura 1 – Fluxograma operacional do guia metodológico para seleção de estudos em revisões da literatura com o auxílio do Claude AI. Alfenas, MG, Brasil, 2026



Fonte: elaborado pelos autores, 2026.

Etapa 1 – Exportação das fontes de informação e primeira deduplicação no EndNote

Após a execução das estratégias de busca nas fontes de informação selecionadas (PubMed/MEDLINE, Embase, LILACS, Scopus, Web of Science e Cochrane Library), todos os registros recuperados devem ser exportados em formato compatível com o gerenciador de referências EndNote. Recomenda-se priorizar arquivos com a extensão .RIS ou o formato nativo de cada base de dados[5].

No EndNote, os registros provenientes das diferentes fontes são consolidados em uma única biblioteca, na qual se realiza a primeira rodada de remoção de duplicatas por meio da função "*Find Duplicates*". Sugere-se configurar o *software* para comparar título, autores e ano de publicação. Além disso, é imperativa a revisão manual do conjunto antes da exclusão definitiva, uma vez que variações tipográficas e de formato — como a presença ou ausência do *Digital Object Identifier* (DOI) e as distintas formas de citação dos autores — podem gerar falsos negativos no algoritmo automático[5,15]. A saída dessa etapa consiste em uma biblioteca consolidada e parcialmente depurada, que deverá ser exportada novamente em formato .RIS para a fase subsequente.

Etapa 2 – Nova deduplicação e fragmentação no Rayyan

Na sequência, a biblioteca é importada para o Rayyan, plataforma inteligente amplamente utilizada na triagem de artigos para revisões. A ferramenta disponibiliza uma versão gratuita, com funcionalidades restritas, e planos pagos, com recursos avançados[6]. O Rayyan dispõe de um algoritmo próprio para a detecção e exclusão de duplicatas, o qual apresenta sensibilidade superior à do EndNote em diversos cenários, justificando a realização de uma segunda rodada de deduplicação nesta etapa[6,15].

Concluída a depuração, o *corpus* consolidado é fracionado em blocos menores, organizados pelo ano de publicação. A fragmentação anual oferece vantagens operacionais relevantes: (i) facilita o rastreamento temporal das publicações e a auditoria do processo de triagem; (ii) gera arquivos de tamanho compatível com a janela de contexto do Claude AI, prevenindo o truncamento de dados; e (iii) possibilita o trabalho simultâneo entre os revisores ou a retomada parcial em casos de interrupção. Cada bloco é exportado nos formatos .CSV ou .RIS e deve conter, no mínimo, os seguintes campos: identificador único, título, autores, ano, periódico, DOI, resumo e palavras-chave.

Etapa 3 – Criação do Projeto no Claude AI e configuração dos critérios de elegibilidade

O recurso de "Projetos" do Claude AI viabiliza a criação de um espaço de trabalho dedicado, no qual arquivos de contexto podem ser anexados de forma persistente e permanecem acessíveis em todas as interações vinculadas àquele ambiente[11]. Essa funcionalidade é o cerne da estratégia aqui proposta, pois transforma os critérios de elegibilidade em um referencial decisório fixo, garantindo consistência durante a avaliação dos diferentes blocos documentais.

Recomenda-se criar um Projeto específico para cada revisão e instruí-lo com dois documentos complementares, anexados como contexto: a) Documento com os critérios de elegibilidade (versão narrativa): contendo os elementos habitualmente descritos no protocolo da revisão, como população, intervenção, comparação, desfechos, desenhos de estudo elegíveis e critérios de exclusão. b) Documento de detalhamento operacional: explicitando as decisões interpretativas para o manejo de casos limítrofes. Pode incluir, por exemplo, a conduta perante estudos com populações mistas, a definição precisa da intervenção, o tratamento de revisões sistemáticas como fontes secundárias, a política referente aos idiomas de publicação e demais convenções acordadas pela equipe.

A segregação entre o documento de critérios e o de detalhamento operacional — estratégia adotada com sucesso na experiência dos autores — apresenta duas vantagens centrais. Primeiro, preserva a integridade do protocolo original (geralmente registrado em plataformas como PROSPERO ou OSF). Segundo, permite o refinamento e o versionamento das regras operacionais ao longo da triagem, sem alterar o documento formalmente registrado. Isso confere flexibilidade ao processo e resguarda a rastreabilidade. A Tabela 1 sintetiza os elementos mínimos recomendados para a composição desses dois documentos.

Tabela 1 – Elementos mínimos recomendados para a configuração dos critérios de elegibilidade no Projeto do Claude AI. Alfenas, MG, Brasil, 2026

Elemento	Conteúdo esperado
População	Definição clara dos participantes, com especificação de idade, sexo, condição clínica e critérios diagnósticos exigidos.
Intervenção	Descrição precisa da intervenção, incluindo definição operacional (e.g., número mínimo de etapas, doses, vias de administração).
Comparação	Grupo de comparação ou padrão de referência adotado (quando aplicável).
Desfechos	Lista dos desfechos primários e secundários considerados elegíveis.
Desenhos elegíveis	Tipos de estudo aceitos (ECR, ECNR, coorte, caso-controle etc.) e desenhos explicitamente excluídos.
Critérios de exclusão	Situações que motivam exclusão imediata, como faixa etária fora do escopo, presença de outras patologias e formatos editoriais.
Idiomas e período	Política sobre idiomas aceitos e janela temporal da busca.
Casos limítrofes	Regras de decisão para populações mistas, dados extraíveis separadamente, revisões prévias, estudos sem resumo, entre outros cenários ambíguos.

Fonte: elaborado pelos autores, 2026.

Passo a passo para a criação do Projeto no Claude AI

Para subsidiar a operacionalização desta etapa, descreve-se a sequência de procedimentos para a criação e a configuração do Projeto no Claude AI, ilustrada na Figura 2.

Figura 2 – Passo a passo para a criação e a configuração de um Projeto no Claude AI, voltado à triagem assistida em revisões da literatura. Alfenas, MG, Brasil, 2026



Fonte: elaborado pelos autores, 2026.

Passo 1 - Acessar o menu lateral e selecionar "*Projects*": Após efetuar o cadastro ou o *login* na interface (https://claude.ai/new) o usuário deve abrir o menu de navegação na lateral esquerda e selecionar a opção "*Projects*" (Figura 2 - Painel A).

Passo 2 - Criar e nomear um novo Projeto: Na tela de listagem, seleciona-se a opção para criar um novo projeto. Recomenda-se utilizar uma nomenclatura facilmente identificável (e.g., a sigla ou o título resumido da revisão) e adicionar uma breve descrição que especifique o tipo de síntese (revisão sistemática, integrativa, de escopo etc.). Essa padronização otimiza a localização e a organização do trabalho em equipe (Figura 2 - Painel B).

Passo 3 - Anexar os arquivos de contexto persistente: No painel direito da interface, na seção "Arquivos", devem-se anexar os dois documentos balizadores da triagem: os critérios de elegibilidade e o detalhamento operacional. Esses arquivos permanecerão disponíveis em todas as interações, assegurando a aplicação consistente dos mesmos parâmetros em cada bloco anual avaliado (Figura 2 - Painel C). Adicionalmente, a interface possibilita a inserção de diretrizes gerais no campo "Instruções", as quais orientarão o comportamento do modelo. O painel "Memória" registrará automaticamente informações relevantes acumuladas durante as interações, fortalecendo a rastreabilidade do método.

Etapa 4 – Triagem assistida por título e resumo no Claude AI

Com o Projeto configurado, inicia-se a execução da triagem. Cada bloco anual de registros é anexado, individualmente, a uma nova conversa dentro do Projeto. O comando (*prompt*) inserido deve solicitar, de forma explí-

cita: (i) a leitura dos títulos e resumos de todos os itens do arquivo; (ii) a aplicação dos critérios de elegibilidade e do detalhamento operacional previamente anexados; e (iii) a geração de uma planilha (.CSV) contendo o identificador do registro, os metadados bibliográficos essenciais, o veredito (incluir para leitura na íntegra ou excluir) e a respectiva justificativa textual.

A justificativa individualizada por registro é o elemento central deste método, pois atua como mecanismo de auditoria e prestação de contas (*accountability*). Ao contrário de sistemas que fornecem apenas classificações binárias, o Claude AI é capaz de articular, em linguagem natural, qual critério exato balizou a decisão (e.g., "revisão narrativa sobre taxanos sem a descrição de um protocolo formal de dessensibilização" ou "estudo pré-clínico realizado em cultura celular"). Essa transparência viabiliza a posterior conferência humana e a mitigação de discordâncias.

Recomenda-se o *download* e o arquivamento das planilhas geradas para cada bloco. Elas atuarão como artefatos do processo de seleção e integrarão o acervo auditável da pesquisa. A Tabela 2 apresenta um modelo de *prompt*, baseado na experiência dos autores e aprimorado iterativamente ao longo da testagem.

Tabela 2 – Modelo de *prompt* utilizado para a triagem de estudos no Claude AI. Alfenas, MG, Brasil, 2026

Solicitação principal:
"Realize a triagem dos estudos disponíveis no arquivo anexo nomeado '[nome_do_bloco].csv' baseado na leitura dos títulos e resumos, conforme os critérios de elegibilidade estabelecidos e disponíveis nos arquivos do projeto nomeados 'critérios de elegibilidade' e 'detalhamento sobre os critérios de elegibilidade'.
Saída esperada:
Produza um novo arquivo .csv contendo, para cada registro analisado, as seguintes colunas:
(1) id_triagem — identificador sequencial atribuído na triagem;
(2) *id_rayyan* (ou *key*) — identificador original do registro no Rayyan;
(3) título do estudo;
(4) ano de publicação;
(5) periódico;
(6) volume, número e páginas;
(7) autores;
(8) DOI;
(9) idioma;
(10) resumo;
(11) palavras-chave;
(12) PubMed ID (quando disponível);
(13) decisão_triagem — 'Incluir para texto completo' ou 'Excluir';
(14) justificativa — explicação textual fundamentada no critério específico de elegibilidade ou exclusão que motivou a decisão.
Instruções adicionais:
• Adote postura conservadora: em caso de dúvida sobre o atendimento aos critérios, marque o registro como 'Incluir para texto completo' e justifique a incerteza.
• Para registros sem resumo, baseie a decisão no título e nas palavras-chave, indicando essa limitação na justificativa.
• Mantenha a codificação *8-bit Unicode Transformation Format* (UTF-8) (com BOM) e o separador vírgula no arquivo .csv gerado, com todos os campos entre aspas para preservar conteúdo com vírgulas e ponto-e-vírgulas".

Fonte: elaborado pelos autores, 2026.

Etapa 5 – Retorno ao Rayyan para conferência e integração ao PRISMA

Após a finalização da triagem pelo Claude AI, é imprescindível reintegrar as decisões ao ambiente do Rayyan. Esse procedimento permite: (i) a execução da conferência humana independente; (ii) a consolidação de todos os dados em um ambiente auditável único; e (iii) a geração automatizada do fluxograma PRISMA, contemplando a contagem precisa das inclusões e exclusões em cada fase metodológica.

Operacionalmente, as planilhas anuais geradas pela IA devem ser compiladas em uma única planilha-mestre. A coluna *id_rayyan* deve ser mantida como chave primária, pois ela possibilita a localização exata de cada registro importado originalmente. Recomenda-se a checagem rigorosa desses identificadores antes do reingresso na plataforma, visando sanar eventuais duplicidades ou falhas de codificação.

No Rayyan, as decisões da IA podem ser aplicadas mediante a atribuição de rótulos (*labels*) personalizados (e.g., 'IA-incluir' e 'IA-excluir'). Essa abordagem preserva o veredito da ferramenta em uma camada informacional distinta daquela gerada pelos pesquisadores. Outra estratégia plausível é a inserção das justificativas geradas pelo modelo no campo de "notas" (*notes*) de cada registro, otimizando a análise comparativa. A escolha do método deve ponderar o volume do *corpus*, o tamanho da equipe de revisores e as preferências do grupo de pesquisa.

A validação humana independente — conduzida por dois revisores em formato duplo-cego no Rayyan — deve avaliar a integralidade das decisões emitidas pela IA. As discordâncias entre o modelo e o pesquisador, ou entre os próprios pesquisadores, devem ser dirimidas por consenso ou mediante o parecer de um terceiro revisor, alinhando-se aos preceitos do *Cochrane Handbook* e do JBI[15,16]. Esse rigor metodológico consagra a IA no papel de primeiro revisor automatizado, mantendo a incontornável supervisão humana, fiadora da qualidade das evidências sintetizadas.

Finalizada a validação, o Rayyan viabiliza a exportação quantitativa dos estudos identificados, deduplicados, inclu-

ídos e excluídos, fornecendo os insumos para a confecção do fluxograma PRISMA. Recomenda-se, por fim, que o relatório final da revisão documente expressamente o uso da IA generativa na etapa de triagem de títulos e resumos. Deve-se incluir uma nota metodológica detalhando os procedimentos de validação adotados, em estrita obediência às diretrizes para o reporte do emprego de inteligência artificial em estudos de revisão[21].

Transparência e registro da experiência de aplicação

A revisão sistemática utilizada como cenário prático para a testagem do fluxo proposto encontra-se registrada na base PROSPERO (CRD420261325732) e está disponível para consulta pública. Os resultados quantitativos completos dessa triagem — incluindo o número de registros identificados, deduplicados, triados pelo Claude AI e confirmados na conferência humana independente, além do respectivo fluxograma PRISMA — serão detalhados na publicação primária do estudo. Essa revisão está atualmente em desenvolvimento pelo mesmo grupo de pesquisa. A opção por não apresentar os dados quantitativos neste momento visa evitar a publicação duplicada de dados e preservar a originalidade da síntese de evidências para o seu artigo principal.

Ressalta-se que o objetivo central deste guia é descrever e fundamentar o fluxo operacional elaborado, e não reportar os achados de uma revisão sistemática específica. Tais resultados serão oportunamente publicizados em um manuscrito próprio, com a devida citação cruzada a este estudo metodológico.

Diante da crescente adoção de ferramentas de IA no apoio à seleção e à triagem de estudos, a Tabela 3 apresenta uma análise comparativa entre a triagem manual, o uso do Claude AI e a aplicação de outras tecnologias. A tabela evidencia as principais diferenças relacionadas aos aspectos operacionais, contextuais e funcionais.

Tabela 3 – Comparação entre triagem manual, Claude AI e outras ferramentas de IA no apoio à seleção de estudos. Alfenas, MG, Brasil, 2026

Aspecto	Triagem manual	Claude AI	Outras IAs (ex.: ChatGPT, Gemini, ASReview)
Tempo médio por bloco (~400 registros)	8 a 15 horas por revisor.	De minutos a poucas horas, conforme o tamanho do arquivo.	Variável. Ferramentas com aprendizado ativo (como o ASReview) reduzem o esforço progressivamente.
Consistência entre lotes	Suscetível à fadiga e à deriva interpretativa ao longo do tempo.	Alta, desde que os arquivos de contexto sejam mantidos inalterados no Projeto.	Depende da memória da sessão. ChatGPT e *Gemine* exigem reconfiguração a cada nova conversa.
Justificativa por registro	Quando registrada, frequentemente é sucinta.	Gerada automaticamente em linguagem natural e vinculada aos critérios de elegibilidade.	O ASReview entrega um escore de probabilidade; ChatGPT e Gemini geram texto livre.
Janela de contexto	Não se aplica.	Extensa; suporta amplos lotes de registros sem truncamento.	ChatGPT e Gemini possuem janelas competitivas; o ASReview não utiliza LLMs.
Critérios persistentes	O revisor consulta o documento do protocolo.	Os arquivos de contexto permanecem fixados ao ambiente do Projeto.	Exige-se colar ou reanexar os critérios manualmente em cada nova interação.
Custo	Tempo de trabalho do pesquisador.	Assinatura mensal; uso direto via interface Claude.ai.	Variável. O ASReview é gratuito; outras soluções comerciais cobram por uso ou assinatura.
Auditabilidade	Alta, caso haja o registro detalhado das decisões pelos revisores.	Alta (gera planilha padronizada contendo o veredito e a justificativa para cada registro).	Variável, a depender da arquitetura e das funcionalidades da ferramenta utilizada.

Fonte: elaborado pelos autores, 2026.

DISCUSSÃO

O presente estudo não teve como propósito o desenvolvimento de uma nova tecnologia de IA, tampouco a configuração de um agente autônomo capaz de operar sem a supervisão humana. O Claude AI foi utilizado em sua forma nativa, por meio de *prompts* direcionados manualmente pelos pesquisadores dentro de um Projeto dedicado, atuando como um modelo de linguagem assistido, e não como um agente de execução independente. A inovação proposta neste estudo reside, portanto, na articulação sistematizada de um fluxo operacional replicável. Esse fluxo integra ferramentas já consolidadas (EndNote e Rayyan) ao uso estruturado de LLMs comerciais, fundamentando-se em um conjunto original de recomendações práticas. Tais recomendações incluem: a separação entre os critérios formais de elegibilidade e o detalhamento operacional para casos limítrofes; a padronização de *prompts* com estrutura de saída auditável; o uso de identificadores únicos (*id_rayyan*) como mecanismo de rastreabilidade entre os ambientes virtuais; e a aplicação de um protocolo de conferência humana independente após a triagem pela IA.

Diferentemente de estudos anteriores — centrados na comparação de desempenho entre a IA e revisores humanos ou em discussões teóricas sobre os riscos da automação[8,12,18]—, o presente guia oferece um modelo teórico-operacional testado empiricamente em uma revisão sistemática em curso. Dessa forma, transforma o uso *ad hoc*

de modelos de linguagem em um processo transparente, auditável e alinhado às diretrizes internacionais de síntese de evidências. Além disso, o fluxo atende às recomendações específicas para o relato transparente do uso da IA como ferramenta metodológica em revisões de literatura[21].

Nesse sentido, o método proposto deve ser analisado sob três prismas interdependentes: como estratégia de eficiência operacional, como intervenção metodológica com riscos inerentes e, fundamentalmente, como vetor de impacto transformador na prática científica em saúde. A incorporação da IA na etapa de seleção de estudos reconfigura o trabalho metodológico das revisões, deslocando o esforço operacional para os sistemas computacionais. Isso libera um maior investimento cognitivo dos revisores para a inferência, a interpretação e a avaliação crítica das decisões analíticas. Essa transformação, no entanto, introduz novas dependências técnicas — como o modelo escolhido, sua versão e seus parâmetros —, além de novas fontes de variação que exigem uma governança rigorosa para manter a rastreabilidade e a responsabilização científica[9].

Potencialidades do Claude AI na otimização da seleção de estudos

No plano da eficiência, o Claude AI demonstra alto potencial ao deslocar o esforço operacional da triagem, por título e resumo, para os sistemas computacionais. Estudos recentes quantificam essa vantagem ao evidenciar que ferramentas baseadas em LLMs podem atingir precisão e revocação comparáveis às de revisores humanos, com redução significativa do tempo necessário para a seleção dos manuscritos. Análises comparativas demonstram *F1-scores* acima de 90% em tarefas de triagem, com a redução do tempo de avaliação de dias para horas na análise de centenas de registros bibliográficos[10,22].

Essa otimização permite a realocação de recursos humanos para etapas de maior exigência cognitiva da pesquisa, como a análise crítica dos achados e a síntese qualitativa. Adicionalmente, a extensa janela de contexto do Claude AI e o recurso de Projetos com arquivos persistentes viabilizam o processamento de amplos lotes de registros sem a fragmentação artificial do trabalho, configurando um diferencial operacional relevante em relação a outras ferramentas de IA[13].

Considerando a crescente adoção dessas tecnologias no apoio à seleção e à triagem de artigos, a Tabela 3 (apresentada na seção de Resultados) propõe uma análise comparativa entre o Claude AI e outros sistemas aplicáveis à mesma finalidade, evidenciando as diferenças relacionadas a aspectos operacionais, contextuais e funcionais. Os dados sintetizados demonstram que, embora plataformas como o ASReview apresentem vantagens específicas em custo e em aprendizado ativo progressivo, o Claude AI destaca-se pela combinação de três fatores: ampla janela de contexto, persistência dos critérios ao longo de toda a triagem e geração de justificativas em linguagem natural vinculadas a cada decisão. Em conjunto, esses atributos favorecem substancialmente a auditabilidade do processo metodológico.

Desafios éticos e metodológicos

O impacto metodológico da IA sobre a validade e a confiabilidade da seleção de estudos exige cautela. Embora a automação reduza falhas humanas decorrentes de fadiga ou inconsistência, ela introduz riscos sistêmicos que demandam gerenciamento rigoroso. Destacam-se os vieses algorítmicos — notadamente os de seleção, de medição e de confundimento —, que podem deslocar o centro de gravidade da síntese para cenários mais compatíveis com o *corpus* de treinamento do modelo. Esse desvio é capaz de gerar assimetrias e classificações conceitualmente frágeis[23-25].

No contexto da seleção de artigos, o viés de seleção pode manifestar-se quando modelos treinados em bases de dados com predomínio de determinados idiomas, delineamentos ou regiões geográficas apresentam desempenho inferior ao avaliar pesquisas de contextos sub-representados, comprometendo a equidade da síntese[23]. Somado a isso, o fenômeno das "alucinações" (geração de informações plausíveis, porém factualmente incorretas) representa um risco epistemológico singular. Em registros sem resumo ou com *abstracts* pouco informativos, a IA pode produzir decisões aparentemente bem fundamentadas, mas embasadas em inferências não verificáveis no texto-fonte[22,26].

A ocorrência de falsos positivos (inclusão indevida de estudos não elegíveis) e de falsos negativos (exclusão indevida de estudos elegíveis) constitui uma preocupação central em qualquer estratégia de triagem assistida por IA, pois impacta diretamente a completude e a validade da síntese final. Na experiência prática que fundamentou o presente guia, esse risco foi mitigado por meio de duas frentes: a orientação explícita para que o modelo adotasse uma postura conservadora perante incertezas (priorizando a inclusão para leitura na íntegra em detrimento da exclusão precoce) e a conferência humana independente da totalidade das decisões geradas. Essa abordagem desloca o risco predominante para os falsos positivos, considerados metodologicamente menos nocivos nesta etapa, visto que serão retificados durante a leitura do texto completo. Os falsos negativos, por outro lado, comprometeriam irreversivelmente o escopo da síntese.

A confiabilidade das decisões, por sua vez, depende intrinsecamente da qualidade e da estabilidade dos *prompt* sutilizados. Comandos genéricos ou inconsistentes entre diferentes execuções podem gerar variabilidade indevida na triagem, mesmo que os critérios de elegibilidade permaneçam imutáveis. A padronização do *prompt* e sua fixação como elemento auditável do processo, conforme proposto neste guia, configuram-se não apenas como boas práticas operacionais, mas como requisitos metodológicos essenciais para garantir a reprodutibilidade da triagem assistida por IA[21].

Para mitigar esses riscos no fluxo estabelecido, recomenda-se a incorporação de mecanismos de contenção. Entre eles, destacam-se a instrução explícita para que o modelo limite suas inferências aos dados contidos no título e no resumo; a exigência de uma justificativa textual atrelada a um critério específico de elegibilidade; e a verificação cruzada por revisores humanos independentes. A transparência e a reprodutibilidade são pilares inegoci-

áveis. Logo, exige-se que a comunidade científica adote práticas de ciência aberta, documentando meticulosamente o modelo de IA, sua versão, o *prompt* e os arquivos de contexto utilizados[23,27].

Segurança, privacidade e transparência no uso da IA

O emprego de ferramentas de IA generativa em processos de pesquisa impõe considerações prementes sobre a segurança e a confidencialidade dos dados, sobretudo ao manipular metadados de produção científica inédita ou protocolos de revisão em andamento. Na experiência de aplicação deste guia, a desativação da configuração "*Help Improve Claude*" na versão Claude Pro configurou-se como medida essencial. O objetivo foi impedir que os registros bibliográficos e os critérios de elegibilidade fossem incorporados aos dados de treinamento de versões futuras da ferramenta[20].

Cumpre ressaltar que essa desativação inibe apenas o uso das conversas para o treinamento do algoritmo, não eliminando o armazenamento dos dados nos servidores da empresa Anthropic durante o período de retenção estipulado em sua política de privacidade. Ademais, a versão de consumidor não oferece as garantias proporcionadas por planos corporativos (como o *Claude for Work*), que contam com contratos rigorosos de proteção, ou por implementações via *Application Programming Interface* (API) configuradas com retenção zero de dados (*Zero Data Retention*). Essa distinção é crucial em revisões que abarcam dados sensíveis, populações vulneráveis ou informações não publicadas. Nesses cenários, a adoção de ambientes com salvaguardas contratuais robustas de confidencialidade é sobremaneira mais adequada.

No fluxo metodológico aqui proposto, a restrição da entrada de dados a metadados bibliográficos não confidenciais (título, resumo, autores, ano, DOI e palavras-chave) reduz a exposição de informações sigilosas. Contudo, não oblitera integralmente o risco, uma vez que resumos de revisões em andamento podem, eventualmente, revelar direcionamentos estratégicos de pesquisas ainda não divulgadas.

Recomenda-se, portanto, que os pesquisadores avaliem rigorosamente a natureza e a sensibilidade dos dados antes da submissão à ferramenta. É imperativo observar a política de privacidade vigente — suscetível a atualizações — e aferir a adequação do plano contratado à natureza do estudo, adotando o nível de proteção contratual compatível com as informações processadas[20].

Limitações do estudo

Apesar de sua relevância, o presente estudo possui limitações metodológicas. A principal delas reside na ausência de validação quantitativa multicêntrica do fluxo proposto. A aplicação prática concentrou-se em um único domínio temático, o que inviabiliza estimar com precisão o desempenho do método sob variações contextuais relevantes, tais como diferentes áreas do conhecimento, terminologias singulares, distintos idiomas de publicação e tipologias variadas de síntese. Outrossim, a eficácia do Claude AI é inerentemente dependente do modelo e da versão empregados, cujas atualizações futuras podem modificar os resultados. Tal dinamicidade impõe a necessidade imperiosa de documentação técnica granular para assegurar a reprodutibilidade metodológica[28].

Implicações para a pesquisa e a prática em saúde

Para a pesquisa em saúde — área marcada por um volume exponencial de publicações e pela premência de respostas rápidas para orientar a prática clínica e as políticas públicas —, a adoção de fluxos automatizados representa um avanço estratégico. O pesquisador, todavia, deve assumir o papel de curador crítico da informação. O Claude AI atua como catalisador da etapa operacional, ao passo que a *expertise* clínica e metodológica do profissional é resguardada para a interpretação de contextos complexos, a apreensão das nuances do cuidado e a avaliação da aplicabilidade das evidências nos serviços de saúde (com ênfase no Sistema Único de Saúde brasileiro, no contexto local).

Nesse novo cenário tecnológico, a validação humana transcende o mero controle de qualidade para consubstanciar-se como parte orgânica do processo metodológico. O papel do revisor é reconfigurado: em vez de realizar a triagem duplicada exaustiva de todo o acervo, sua atenção converge para a reconciliação de discrepâncias entre as decisões da IA e a avaliação humana inicial. O esforço cognitivo é, então, direcionado para a análise crítica e para a interpretação contextualizada das evidências[22].

Adicionalmente, o processo de elaboração e configuração dos critérios de elegibilidade no Projeto fomenta uma reflexão profunda sobre o protocolo de revisão. Frequentemente, essa etapa revela ambiguidades operacionais que, de outro modo, emergiriam apenas no decurso da triagem. Esse atributo confere ao método uma dimensão pedagógica substancial, de particular valia na formação continuada de pesquisadores em síntese de evidências.

Direções futuras

As futuras investigações devem concentrar esforços no desenvolvimento de diretrizes metodológicas específicas para o emprego da IA na seleção de estudos, estabelecendo padrões mínimos de reporte (especificação do modelo, versão, *prompt*, parâmetros e trilhas de auditoria). Impõe-se também a condução de estudos de validação externa multicêntrica. Tais pesquisas devem confrontar o desempenho da IA com o de revisores humanos em domínios temáticos heterogêneos, valendo-se de métricas precisas de sensibilidade e especificidade, além de avaliar o impacto concreto da ferramenta nas conclusões finais da síntese[28,29]. A sedimentação dessas diretrizes instituirá os alicerces necessários para o uso responsável, ético e auditável da IA generativa na produção de evidências em saúde.

CONCLUSÃO

A incorporação do Claude AI — um LLM desenvolvido pela empresa Anthropic — na etapa de seleção de estudos em revisões de literatura moderniza o processo de síntese de evidências, uma vez que automatiza uma das fases mais morosas da produção científica: a triagem por título e resumo. Este estudo, de natureza teórico-metodológica, apresenta uma proposta de fluxo estruturado em cinco eta-

pas para orientar o uso responsável dessa tecnologia. A abordagem integra o rigor científico aos recursos de Processamento de Linguagem Natural (PLN), demonstrando que a automação pode ser adotada com segurança mediante o estabelecimento de protocolos claros, supervisão humana incontornável e validação contínua.

A despeito da elevada capacidade do Claude AI para interpretar a linguagem científica e aplicar critérios de elegibilidade — quando configurados como contexto persistente em um Projeto dedicado —, sua utilização demanda cautela em virtude dos riscos inerentes de vieses algorítmicos e "alucinações". Essa ressalva é particularmente crítica em revisões destinadas a fundamentar decisões clínicas ou a balizar políticas de saúde. Portanto, o uso ético e responsável da ferramenta impõe a garantia de transparência, rastreabilidade e reprodutibilidade em todas as etapas metodológicas.

Em termos de implicações práticas, recomenda-se que os pesquisadores: (1) configurem o Projeto no Claude AI com arquivos de contexto persistentes, englobando os critérios de elegibilidade e o seu respectivo detalhamento operacional, a fim de assegurar a consistência analítica entre os blocos de triagem; (2) utilizem comandos (*prompts*) estruturados e validados iterativamente, exigindo explicitamente uma justificativa textual para cada registro avaliado; (3) realizem a deduplicação prévia nos *softwares* EndNote e Rayyan, fragmentando o *corpus* bibliográfico em blocos anuais para otimizar o processamento e viabilizar a auditoria; e (4) preservem a validação humana independente (em duplicata) sobre as decisões geradas pela Inteligência Artificial, solucionando eventuais divergências por consenso ou por intermédio do parecer de um terceiro revisor.

CONFLITO DE INTERESSES

Os autores declaram não haver conflito de interesses.

USO DE INTELIGÊNCIA ARTIFICIAL

O Claude AI foi utilizado como suporte metodológico e operacional durante a elaboração do artigo, incluindo auxílio na organização do conteúdo e na geração das imagens apresentadas no manuscrito. O ChatGPT foi empregado como ferramenta auxiliar para revisão ortográfica, aprimoramento da fluidez textual e refinamento da redação científica. Ademais, todas as decisões metodológicas, análises, interpretações, organização dos dados e conteúdo científico do estudo foram integralmente realizados e validados pelos autores.

REFERÊNCIAS

1. Farhat N, Tsaioun K, Saunders-Hastings P, Morgan RL, Ramoju S, Hartung T, et al. Systematic review in evidence-based risk assessment. ALTEX. 2022;39(3):463-79. https://doi.org/10.14573/altex.2004111. PMID: 34585732.
2. Wu Q, Nabi G. Systematic reviews and data analyses: guide to clinical practice and research. Scott Med J. 2023;68(3):69-71. https://doi.org/10.1177/00369330231193888. PMID: 37605529.
3. Canto GDL. Revisões sistemáticas da literatura: guia prático. Curitiba: Brazil Publishing; 2020.
4. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS Med. 2010;7(9):e1000326. https://doi.org/10.1371/journal.pmed.1000326. PMID: 20877712.
5. Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open. 2017;7(2):e012545. https://doi.org/10.1136/bmjopen-2016-012545. PMID: 28242767.
6. Bramer WM, Giustini D, Kramer BMR, Anderson PF. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. Syst Rev. 2013;2:115. https://doi.org/10.1186/2046-4053-2-115. PMID: 24360284.
7. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan - a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210. https://doi.org/10.1186/s13643-016-0384-4. PMID: 27919275.
8. Ge L, Agrawal R, Singer M, Kannapiran P, Molina JADC, Teow KL, et al. Leveraging artificial intelligence to enhance systematic reviews in health research: advanced tools and challenges. Syst Rev. 2024;13(1):269. https://doi.org/10.1186/s13643-024-02682-2. PMID: 39456077.
9. Maity S, Saikia MJ. Large language models in healthcare and medical applications: a review. Bioengineering (Basel). 2025;12(6):631. https://doi.org/10.3390/bioengineering12060631. PMID: 40564447.
10. Guo E, Gupta M, Deng J, Park YJ, Paget M, Naugler C. Automated paper screening for clinical reviews using large language models: data analysis study. J Med Internet Res. 2024;26:e48996. https://doi.org/10.2196/48996. PMID: 38214966.
11. Syriani E, David I, Kumar G. Screening articles for systematic reviews with ChatGPT. J Comput Lang. 2024;80:101287. https://doi.org/10.1016/j.cola.2024.101287
12. van Dijk SHB, Brusse-Keizer MGJ, Bucsán CC, van der Palen J, Doggen CJM, Lenferink A. Artificial intelligence in systematic reviews: promising when appropriately used. BMJ Open. 2023;13(7):e072254. https://doi.org/10.1136/bmjopen-2023-072254. PMID: 37419641.
13. Claude [Internet]. San Francisco: Anthropic; 2026 [citado 2026 Maio 18]. Disponível em: https://claude.ai/
14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71. PMID: 33782057.

15. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. Version 6.5 (updated August 2024). London: Cochrane; 2024 [citado 2026 Maio 18]. Disponível em: https://www.cochrane.org/handbook
16. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis [Internet]. Adelaide: JBI; 2024 [citado 2026 Maio 18]. Disponível em: https://synthesismanual.jbi.global/. https://doi.org/10.46658/JBIMES-24-01
17. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs. 2005;52(5):546-53. https://doi.org/10.1111/j.1365-2648.2005.03621.x. PMID: 16268861.
18. Wang S, Scells H, Koopman B, Zuccon G. Can ChatGPT write a good Boolean query for systematic review literature search? [Internet]. In: Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23); 2023 Jul 23-27; Taipei, Taiwan. New York (NY): ACM; 2023 [citado 2026 Maio 18]. p. 1426-36. Disponível em: https://dl.acm.org/doi/10.1145/3539618.3591703. https://doi.org/10.1145/3539618.3591703
19. Echer IC. Elaboração de manuais de orientação para o cuidado em saúde. Rev Lat Am Enfermagem. 2005;13(5):754-7. https://doi.org/10.1590/S0104-11692005000500022. PMID: 16308635.
20. How do I change my model improvement privacy settings? [Internet]. San Francisco: Anthropic; 2026 [citado 2026 Ago 23]. Disponível em: https://privacy.claude.com/en/articles/12109829-how-do-i-change-my-model-improvement-privacy-settings
21. Holst D, Moenck K, Koch J, Schmedemann O, Schüppstuhl T. Transparent reporting of AI in systematic literature reviews: development of the PRISMA-trAIce checklist. JMIR AI. 2025;4:e80247. https://doi.org/10.2196/80247. PMID: 41370833.
22. Khraisha Q, Put S, Kappenberg J, Warraitch A, Hadfield K. Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. Res Synth Methods. 2024;15(4):616-26. https://doi.org/10.1002/jrsm.1715. PMID: 38484744.
23. Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, et al. Fairness of artificial intelligence in healthcare: review and recommendations. Jpn J Radiol. 2024;42(1):3-15. https://doi.org/10.1007/s11604-023-01474-3. PMID: 37540463.
24. Nazer LH, Zatarah R, Waldrip S, Chen Ke JX, Moukheiber M, Khanna AK, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. PLOS Digit Health. 2023;2(6):e0000278. https://doi.org/10.1371/journal.pdig.0000278. PMID: 37347721.
25. Chen F, Wang L, Hong J, Jiang J, Zhou L. Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models. J Am Med Inform Assoc. 2024;31(5):1172-83. https://doi.org/10.1093/jamia/ocae060. PMID: 38520723.
26. Dhuliawala S, Komeili M, Xu J, Raileanu R, Li X, Celikyilmaz A, et al. Chain-of-verification reduces hallucination in large language models [Internet]. In: Findings of the Association for Computational Linguistics: ACL 2024. Stroudsburg: Association for Computational Linguistics; 2024 [citado 2026 Ago 23]. p. 3563-78. Disponível em: https://aclanthology.org/2024.findings-acl.212/. https://doi.org/10.18653/v1/2024.findings-acl.212
27. Ioannidis JP. Transparency, bias, and reproducibility across science: a meta-research view. J Clin Invest. 2024;134(22):e181923. https://doi.org/10.1172/JCI181923. PMID: 39545412.
28. Siontis GCM, Sweda R, Noseworthy PA, Friedman PA, Siontis KC, Patel CJ. Development and validation pathways of artificial intelligence tools evaluated in randomised clinical trials. BMJ Health Care Inform. 2021;28(1):e100466. https://doi.org/10.1136/bmjhci-2021-100466. PMID: 34969668.
29. Yuan H. Toward real-world deployment of machine learning for health care: external validation, continual monitoring, and randomized clinical trials. Health Care Sci. 2024;3(5):360-64. https://doi.org/10.1002/hcs2.114. PMID: 39479276.

CONTRIBUIÇÃO DE AUTORIA

Concepção do estudo: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Obtenção de dados: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Análise de dados: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Interpretação dos dados: Sampaio RS, Fracaroli YR, Franco APMML, Rodrigues MA, Nascimento FC, Lima CS, Costa ICPS.
Todos os autores se responsabilizam pela redação textual e revisão crítica do conteúdo intelectual, pela versão final publicada e por todos os aspectos éticos, legais e científicos relacionados à exatidão e à integridade do estudo.

© 2026 Autores. Os autores mantêm os direitos autorais e concedem à Online Brazilian Journal of Nursing o direito de primeira publicação. Este artigo está licenciado sob a licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0).