



INTELIGÊNCIA ARTIFICIAL GENERATIVA E MODELOS DE LINGUAGEM DE GRANDE ESCALA COMO FERRAMENTAS DE APOIO METODOLÓGICO NA PESQUISA EM SAÚDE: UMA ANÁLISE BIBLIOMÉTRICA

GENERATIVE ARTIFICIAL INTELLIGENCE AND LARGE LANGUAGE MODELS AS METHODOLOGICAL SUPPORT TOOLS IN HEALTH RESEARCH: A BIBLIOMETRIC ANALYSIS

INTELENCIA ARTIFICIAL GENERATIVA Y MODELOS DE LENGUAJE DE GRAN TAMAÑO COMO HERRAMIENTAS DE APOYO METODOLÓGICO EN LA INVESTIGACIÓN EN SALUD: UN ANÁLISIS BIBLIOMÉTRICO

Yasmin Saba de Almeida¹
Elaine Antunes Cortez¹
Geilsa Soraia Cavalcanti Valente¹

ORCID: 0000-0003-2391-7009
ORCID: 0000-0003-3912-9648
ORCID: 0000-0003-4488-4912

¹ Universidade Federal Fluminense, Escola de Enfermagem Aurora de Afonso Costa. Niterói, Rio de Janeiro, Brasil.

Como citar: Almeida YS, Cortez EA, Valente GSC. Generative Artificial Intelligence and Large Language Models as methodological support tools in health research: a bibliometric analysis. Online Braz J Nurs. 2026;25(Suppl1):e20267204. <https://doi.org/10.17665/1676-4285.20267204>

RESUMO

Objetivo: Mapear a produção científica e as tendências temáticas do uso de inteligência artificial generativa (GenAI) e modelos de linguagem de grande escala (LLMs) como ferramentas de apoio metodológico na pesquisa em saúde. **Método:** Estudo bibliométrico realizado em setembro de 2026 nas bases PubMed, Scopus e Web of Science. Os registros foram triados por meio de um pipeline híbrido (humano-IA) com apoio de um LLM segundo critérios previamente definidos e revisão humana. Os dados foram analisados no RStudio com uso do pacote *bibliometrix* e de sua interface *biblioshiny*. **Resultados:** O corpus final compreendeu 552 documentos, distribuídos em 337 fontes, com idade média de 0,88 anos e taxa anual de crescimento de 70,7%. Os Estados Unidos concentraram o maior número de afiliações de autores, seguidos pela China. A literatura apresentou alta dispersão entre periódicos e forte concentração em autores ocasionais (89,2% com apenas uma publicação), além de modesta colaboração internacional (14,7%). O mapa temático revelou que os temas motores permanecem centrados no uso genérico de LLMs e ChatGPT, enquanto aplicações metodológicas específicas ainda não constituem agrupamentos consolidados. **Conclusão:** O uso da GenAI e dos LLMs como apoio metodológico à pesquisa em saúde constitui um campo emergente e em rápida expansão, mas ainda desprovido de uma estrutura temática e de canais de publicação consolidados. A consolidação da área exige o estabelecimento de critérios normativos de validação, transparência no relato, supervisão humana estrita e letramento dos pesquisadores em inteligência artificial, garantindo a integridade das evidências científicas produzidas, com repercussões diretas na qualificação das práticas e na segurança do cuidado em saúde.

Descritores: Bibliometria; Inteligência Artificial Generativa; Modelos de Linguagem de Grande Escala; Saúde; Técnicas de Pesquisa.

ABSTRACT

Objective: To map the scientific output and thematic trends in the use of generative artificial intelligence (GenAI) and large language models (LLMs) as methodological support tools in health research. **Method:** Bibliometric study conducted in September 2026 in the PubMed, Scopus, and Web of Science databases. Records were screened through a hybrid (human-AI) pipeline, in which an LLM applied predefined criteria and a human reviewer checked the results. Data were analyzed in RStudio using the *bibliometrix* package and its *biblioshiny* interface. **Results:** The final corpus comprised 552 documents from 337 sources, with a mean document age of 0.88 years and an annual growth rate of 70.7%. The United States accounted for the largest number of author affiliations, followed by China. The literature showed high dispersion across journals and a strong concentration of occasional authors (89.2% with a single publication), along with modest international collaboration (14.7%). The thematic map revealed that motor themes remain centered on the generic use of LLMs and ChatGPT, whereas specific methodological applications have not yet formed consolidated clusters. **Conclusion:** The use of GenAI and LLMs as methodological support in health research is an emerging and rapidly expanding field, but it still lacks a consolidated thematic structure and established publication venues. Consolidating the field will require normative validation criteria, transparent reporting, strict human oversight, and AI literacy among researchers, so as to safeguard the integrity of the scientific evidence produced, with direct implications for the quality of practice and the safety of health care.

Descriptors: Bibliometrics; Generative Artificial Intelligence; Large Language Models; Health; Investigative Techniques.

RESUMEN

Objetivo: Mapear la producción científica y las tendencias temáticas del uso de la inteligencia artificial generativa (GenAI) y de los modelos de lenguaje de gran tamaño (LLM) como herramientas de apoyo metodológico en la investigación en salud. **Método:** Estudio bibliométrico realizado en septiembre de 2026 en las bases de datos PubMed, Scopus y Web of Science. Los registros se cribaron mediante un flujo híbrido (humano-IA) en el que un LLM aplicó criterios definidos previamente y una revisora humana verificó los resultados. Los datos se analizaron en RStudio con el paquete *bibliometrix* y su interfaz *biblioshiny*. **Resultados:** El corpus final estuvo compuesto por 552 documentos, distribuidos en 337 fuentes, con una antigüedad media de 0,88 años y una tasa de crecimiento anual del 70,7%. Estados Unidos concentró el mayor número de afiliaciones de autores, seguido de China. La literatura mostró una alta dispersión entre revistas y una fuerte concentración de autores ocasionales (el 89,2% con una sola publicación), además de una colaboración internacional modesta (14,7%). El mapa temático reveló que los temas motores siguen centrados en el uso genérico de los LLM y de ChatGPT, mientras que las aplicaciones metodológicas específicas aún no conforman agrupaciones consolidadas. **Conclusión:** El uso de la GenAI y de los LLM como apoyo metodológico a la investigación en salud constituye un campo emergente y en rápida expansión, pero que todavía carece de una estructura temática y de canales de publicación consolidados. La consolidación del área exige establecer criterios normativos de validación, transparencia en la comunicación de los resultados, una supervisión humana estricta y la alfabetización de los investigadores en inteligencia artificial, de modo que se garantice la integridad de la evidencia científica producida, con repercusiones directas en la calidad de las prácticas y en la seguridad de la atención en salud.

Descriptores: Bibliometría; Inteligencia Artificial Generativa; Modelos de Lenguaje de Gran Escala; Salud; Técnicas de Investigación.

Editores:

Rosimere Ferreira Santana (ORCID: 0000-0002-4593-3715)
Audrey Vidal Pereira (ORCID: 0000-0002-6570-9016)

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

Autor correspondente:

Yasmin Saba de Almeida
E-mail: yasminsaba@id.uff.br

O que já se sabe:

- A inteligência artificial generativa e os modelos de linguagem de grande escala ampliaram o uso da inteligência artificial para atividades de criação, análise e organização do conhecimento.
- Na saúde, essas tecnologias são empregadas no apoio diagnóstico, na educação, na redação acadêmica e na descoberta e síntese da literatura científica.
- O uso da inteligência artificial generativa como apoio metodológico às etapas da pesquisa em saúde ainda é pouco explorado de forma sistemática.

O que este artigo acrescenta:

- Apresenta um panorama bibliométrico específico do uso da inteligência artificial generativa como apoio metodológico à pesquisa em saúde.
- Identifica os principais autores, periódicos, países, redes de colaboração e tendências temáticas desse campo.
- Evidencia um campo recente e em rápida expansão e indica prioridades para sua consolidação científica e metodológica.

INTRODUÇÃO

A inteligência artificial generativa (GenAI), especialmente os modelos de linguagem de grande escala (LLMs), transicionou rapidamente de um recurso de redação assistida para aplicação em domínios de alta complexidade epistêmica, incluindo diferentes etapas estruturantes do processo de pesquisa científica⁽¹⁾. No ecossistema biomédico, essa mudança paradigmática foi acompanhada pela expansão de ferramentas assistidas por inteligência artificial (IA) voltadas à descoberta e exploração da literatura, que incluem desde plataformas de mineração textual até a representação relacional e semântica do conhecimento^(2,3).

No âmbito da pesquisa em saúde, a síntese de evidências emergiu como um dos principais domínios de aplicação da GenAI e dos LLMs. Esses modelos vêm sendo progressivamente incorporados à formulação de estratégias de busca, à triagem automatizada ou semiautomatizada de literatura, à extração parametrizada de dados e avaliação da qualidade metodológica ou do risco de viés. No entanto, evidências sintetizadas por revisão de escopo recente advertem que a maturidade tecnológica dessas ferramentas ainda não assegura sua reprodutibilidade ou acurácia isenta de supervisão crítica⁽⁴⁾.

Embora o volume de publicações sobre IA em saúde tenha apresentado crescimento exponencial, as investigações dedicadas às aplicações estritamente metodológicas no fluxo de pesquisa permanecem diluídas entre estudos focados em auxílio diagnóstico, educação em saúde e suporte à escrita acadêmica. Buscas preliminares nas bases PubMed, Scopus e Web of Science confirmaram a ausência de estudos bibliométricos que mapeiem, com rigor científico e detalhamento sistemático, o uso dessas tecnologias como apoio metodológico na pesquisa em saúde. Essa lacuna é relevante porque a pesquisa em saúde depende da produção oportuna e rigorosa de evidências para orientar o cuidado, a elaboração de diretrizes e as políticas de saúde. Ao apoiar etapas que exigem grande volume de trabalho, como busca, triagem, extração e síntese de dados, a GenAI e os LLMs têm elevado potencial para otimizar o ciclo de produção de evidências, desde que sua validação técnica, acurácia e reprodutibilidade procedimental sejam devidamente asseguradas^(4,5).

Nesse sentido, compreender como esse campo vem se constituindo permite reconhecer quais etapas da pesquisa em saúde já concentram esforços, quais autores, instituições e fontes impulsionam a produção científica e quais lacunas ainda permanecem. Essa visão pode orientar pri-

oridades de pesquisa e favorecer o uso da GenAI e dos LLMs de forma mais consistente com as necessidades e o rigor da pesquisa em saúde. A análise bibliométrica permite examinar o crescimento da produção, as redes de colaboração e os temas que estruturam o conhecimento produzido, oferecendo um panorama preciso do estágio de maturidade desse campo^(6,7). Assim, este estudo tem como objetivo mapear a produção científica e as tendências temáticas do uso de GenAI e LLMs como ferramentas de apoio metodológico na pesquisa em saúde.

MÉTODO

Trata-se de um estudo cientométrico e bibliométrico, caracterizado pela aplicação de métodos quantitativos para analisar a produção científica em um campo específico do conhecimento. A bibliometria utiliza indicadores estatísticos para examinar padrões de publicação, citação e colaboração científica, permitindo identificar estruturas do conhecimento, a dinâmica de produtividade e os domínios temáticos de uma disciplina^(6,7).

Questão de pesquisa

A questão norteadora do estudo foi: qual é o perfil bibliométrico e quais são as tendências temáticas da produção científica sobre GenAI e LLMs utilizados como ferramentas de apoio metodológico na pesquisa em saúde?

Critérios de elegibilidade

Foram incluídos documentos nos quais a GenAI, especialmente os LLMs, fosse empregada, avaliada ou discutida especificamente como apoio a pelo menos uma etapa metodológica da pesquisa em saúde. Foram consideradas etapas como concepção e delineamento metodológico, elaboração de estratégias de busca, triagem de títulos, resumos ou textos completos, extração e síntese de dados, avaliação da qualidade metodológica ou do risco de viés, análise quantitativa ou qualitativa de dados e geração de hipóteses de pesquisa.

Foram excluídos documentos que abordavam essas tecnologias exclusivamente em contextos de prática clínica ou assistencial, educação ou capacitação em saúde, redação científica ou edição de manuscritos, ou que apenas mencionavam IA sem aplicação metodológica identificável. Também foram excluídos estudos fora do campo da saúde.

Não foram aplicadas restrições quanto ao idioma, tipo

documental ou período de publicação.

Busca bibliográfica

As buscas foram realizadas em 1º de setembro de 2026 nas bases de dados indexadas PubMed, Scopus e Web of Science. As estratégias combinaram o vocabulário controlado do *Medical Subject Headings* (MeSH) com termos livres relacionados à GenAI e aos LLMs, ao processo de pesquisa e à pesquisa em saúde. Também foram utilizados operadores booleanos e termos de exclusão para reduzir a recuperação de estudos voltados exclusivamente a aplicações clínicas, educação em saúde ou redação científica. As estratégias completas de busca estão disponíveis e auditáveis no repositório do projeto no Open Science Framework (OSF), sob DOI: 10.17605/OSF.IO/NC8K9.

Seleção de estudos assistida por GenAI

Foi desenvolvido, em Python, um fluxo computacional para automatizar e registrar a deduplicação, a classificação dos registros e o registro das decisões de triagem. O fluxo foi construído com apoio das ferramentas de GenAI Codex, da OpenAI, e Claude Code, da Anthropic, utilizando o Visual Studio Code, versão 1.137.0, como ambiente de desenvolvimento. Os critérios de elegibilidade e o *prompt* estruturado foram definidos e refinados iterativamente pela pesquisadora ao longo desse processo. O código-fonte e a versão do *prompt* utilizada na classificação estão disponíveis no repositório público enfermagempesquisadora/genai-health-research-screening no GitHub.

Após a deduplicação por correspondência exata e semântica, os títulos e resumos foram processados com apoio do LLM Claude. Para cada registro, o modelo indicou a decisão de incluir, excluir ou classificar como incerto, além de registrar a categoria de exclusão, quando aplicável, uma justificativa concisa, um trecho do título ou resumo que sustentasse a decisão e o nível de confiança informado pelo modelo. Os trechos de suporte foram verificados automaticamente em relação ao texto-fonte. A pesquisadora realizou conferência geral das classificações e avaliou prioritariamente os registros classificados como incertos, definindo as decisões finais de triagem sob supervisão humana integral.

Processamento e análise dos dados

Os dados foram extraídos, harmonizados e analisados no ambiente R (via interface RStudio), com o pacote *bibliometrix*⁽⁶⁾ e sua interface gráfica, o *biblioshiny*. O *bibliometrix* é uma ferramenta de código aberto voltada à análise bibliométrica, com recursos avançados para avaliação de produtividade, colaboração, citação e mapeamento científico⁽⁶⁾. Análises descritivas complementares foram realizadas em planilhas do Microsoft Excel.

A análise informétrica foi estruturada com base nas três leis clássicas da bibliometria: 1) Lei de Bradford (avaliou a dispersão dos artigos entre as fontes de publicação, permitindo a delimitação empírica do núcleo de periódicos do corpus)⁽⁸⁾; 2) Lei de Lotka (examinou a distribuição da produtividade dos autores, testando o ajuste estatístico entre a distribuição observada e o modelo teórico de potência, via

teste de Kolmogorov-Smirnov)⁽⁹⁾; 3) Lei de Zipf (descreveu a distribuição da frequência dos termos, fundamentando as análises de coocorrência de palavras-chave e a estrutura conceitual do campo)⁽¹⁰⁾.

Também foram analisados o volume anual de produção, a taxa anual de crescimento, o tipo e o idioma dos documentos, a idade média das publicações, a distribuição geográfica das afiliações dos autores, as instituições mais representadas, os países mais produtivos e citados, as fontes e os autores mais produtivos ou citados, a coocorrência de palavras-chave e o mapeamento temático. Os indicadores de produção e impacto incluíram número de publicações, citações totais e citações por documento. As citações foram consideradas conforme os valores presentes nos metadados dos registros importados e interpretadas como medidas descritivas do corpus, considerando as diferenças de cobertura entre as bases⁽¹¹⁾.

O diagrama estratégico posicionou os agrupamentos temáticos em um espaço bidimensional conforme a centralidade (grau de relevância e conectividade intergrupos) e a densidade (grau de desenvolvimento interno do agrupamento). Os temas foram classificados em quatro quadrantes: motores, de nicho, emergentes ou em declínio e básicos ou transversais.

Aspectos éticos

Por utilizar exclusivamente registros bibliográficos e informações de acesso público, sem dados identificáveis de participantes, o estudo não demandou apreciação pelo Sistema CEP/Conep, conforme a Resolução CNS nº 510/2016⁽¹²⁾. Ferramentas de GenAI foram utilizadas como apoio ao desenvolvimento do fluxo de análise, à triagem dos registros e à redação do manuscrito, sob supervisão da pesquisadora e responsabilidade intelectual direta das pesquisadoras.

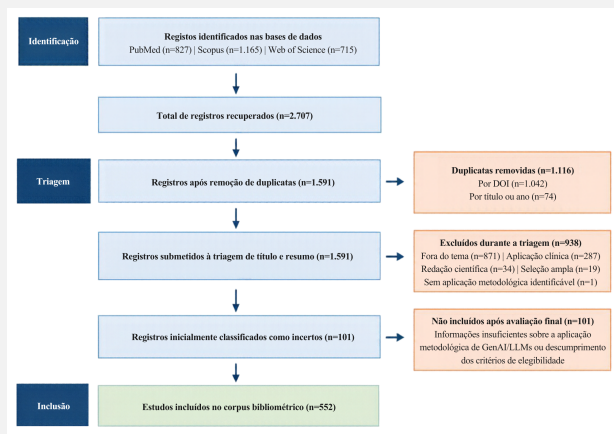
RESULTADOS

Fluxo de seleção e elegibilidade do corpus

Foram identificados 2.707 registros nas bases selecionadas. Após a remoção de 1.116 duplicatas, sendo 1.042 identificadas por DOI e 74 por título e ano de publicação, 1.591 registros únicos foram submetidos à triagem de títulos e resumos. O pipeline híbrido de triagem (humano-IA) resultou em 552 registros elegíveis para o corpus bibliométrico e 938 excluídos. Entre os registros excluídos, 587 estavam fora do tema, 287 abordavam aplicações clínicas, 34 tratavam de redação científica, 29 de educação em saúde e um não apresentava aplicação metodológica própria identificável.

Os 101 registros inicialmente classificados como incertos foram excluídos do corpus final após avaliação qualitativa, pois não foi possível confirmar que a GenAI ou os LLMs fossem utilizados, avaliados ou apoiassem uma etapa metodológica da pesquisa em saúde. Entre eles, havia também estudos fora do tema ou voltados a aplicações clínicas, redação científica, educação em saúde e documentos sem conteúdo metodológico próprio. A Figura 1 apresenta o fluxo de seleção dos estudos.

Figura 1 – Fluxograma de seleção dos estudos. Niterói, RJ, Brasil, 2026



Fonte: elaborado pelos autores, 2026.

Características e evolução da produção científica

Os 552 documentos foram distribuídos em 337 fontes. Predominaram artigos originais de periódicos (n=425; 77,0% do corpus), seguidos por trabalhos em anais de congressos (n=68), revisões (n=38), cartas (n=13), editoriais (n=3), capítulos de livro (n=3) e notas (n=2). O idioma inglês foi predominantemente empregado, estando presente em 98,7% dos registros. A taxa média de citações por documento foi de 2,62.

A produção científica abrangeu o período bibliográfico de 2016 a 2027 (Figura 2). Foi identificado um documento em 2016, sem registros entre 2017 e 2022. A partir de 2023, observou-se aumento expressivo da produção, com 19 documentos em 2023, 98 em 2024, 223 em 2025 e 210 documentos até 1º de setembro de 2026. O único documento registrado em 2027 correspondia a um trabalho de evento já disponível online em 2026, mas vinculado a volume bibliográfico de 2027. O registro foi mantido no corpus e não foi interpretado como tendência de produção posterior ao encerramento das buscas. A idade média dos documentos foi de 0,88 anos, evidenciando a extrema contemporaneidade do campo.

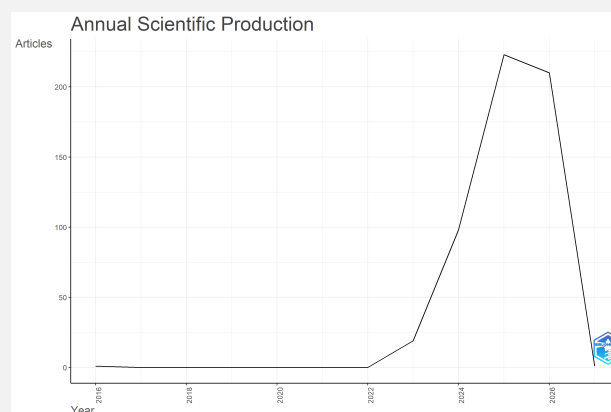
Dispersão das fontes de publicação e impacto de citação (Lei de Bradford)

A distribuição das 337 fontes em zonas de Bradford identificou um núcleo de 27 fontes, responsável por 187 documentos, correspondentes a 33,9% do corpus (Figura 3). A segunda zona reuniu 127 fontes e 182 documentos, enquanto a terceira incluiu 183 fontes e 183 documentos. O modelo de Bradford apresentou excelente ajuste linear ($R^2 = 0,9249$) com um multiplicador de 3,07. O teste de Kolmogorov-Smirnov confirmou a descontinuidade estatística entre a distribuição empírica observada e o modelo teórico ideal ($D = 0,1899$; $p < 0,001$).

O periódico *Studies in Health Technology and Informatics* destacou-se como a fonte mais produtiva (n=26). Em termos de impacto citacional, o *Journal of Clinical Epidemiology* liderou o corpus com 149 citações acumuladas

em seis publicações (Tabela 1).

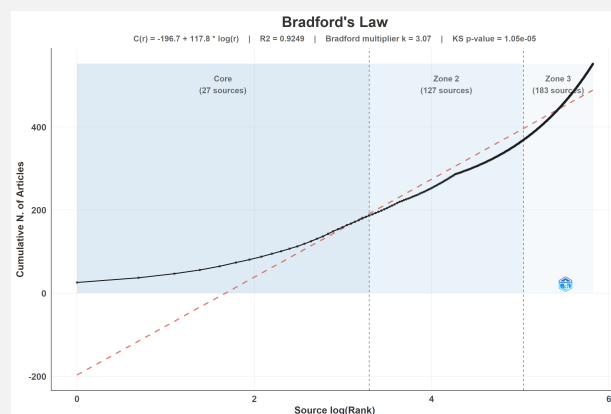
Figura 2 – Produção científica anual sobre GenAI e LLMs como ferramentas de apoio metodológico na pesquisa em saúde, 2016–2027 (n=552). Niterói, RJ, Brasil, 2026



Nota: O único documento com ano bibliográfico de 2027 estava disponível online em 2026 e foi mantido no corpus por ter sido recuperado durante o período de busca.

Fonte: dados da pesquisa, gerados no Biblioshiny, 2026.

Figura 3 – Distribuição das fontes em zonas de Bradford. Niterói, RJ, Brasil, 2026



Fonte: dados da pesquisa, gerados no Biblioshiny, 2026.

Entre os artigos com maior contagem global de citações, destacaram-se os voltados à exploração semântica e síntese de evidências. O artigo de Wei CH, publicado em 2024 abordando a plataforma PubTator 3.0, foi o mais citado do corpus (n=104), seguido pelos estudos de Lieberum JL (2025; n=103), Aljamaan F (2024; n=97), Wang S (2024; n=92) e Jin Q (2024; n=91).

Padrões de autoria, produtividade e cooperação (Lei de Lotka)

O corpus reuniu 2.914 autores. Wang Y foi o autor mais produtivo (n=18), seguido por Wang Z (n=13), Chen Y (n=12) e Li Y (n=12). A produção foi predominantemente colaborativa, com média de 6,32 coautores por artigo. Apenas 21 documentos, correspondentes a 3,8%, foram publicados por autoria individual. O índice de colaboração internacional foi de 14,7%.

Tabela 1 – Principais fontes do corpus segundo número de publicações e citações disponíveis. Niterói, RJ, Brasil, 2026

Fonte	NP	TC	CPP (TC/NP)	% do total (NP/552 × 100)
Studies in Health Technology and Informatics	26	11	0,42	4,71
Journal of Medical Internet Research	11	0	0,00	1,99
bioRxiv: The Preprint Server for Biology	10	0	0,00	1,81
JMIR AI	9	0	0,00	1,63
Journal of the American Medical Informatics Association	9	0	0,00	1,63
Scientific Reports	9	0	0,00	1,63
Bioinformatics (Oxford, England)	7	0	0,00	1,27
BMC Medical Research Methodology	7	33	4,71	1,27
Cureus	7	0	0,00	1,27
BMJ Open	6	1	0,17	1,09

Nota: NP = número de publicações; TC = total de citações; CPP = citações por publicação.

Fonte: dados da pesquisa, 2026.

A aplicação da Lei de Lotka revelou que 2.599 autores (89,2% do total) publicaram apenas um artigo na área (Tabela 2). A distribuição empírica diferiu do modelo teórico clássico da Lei de Lotka, com expoente $\beta = 2,0$ ($D = 0,6429$; $p = 0,006$). No entanto, a estimativa do modelo de potência ajustado ao corpus apresentou $\beta = 2,87$ (constante $C =$

0,444 e $R^2 = 0,928$), não apresentando diferença estatisticamente significativa em relação aos dados observados ($D = 0,2857$; $p = 0,617$). Esse elevado expoente β reflete uma concentração atípica de autores ocasionais, característica de campos emergentes.

Tabela 2 – Distribuição dos autores segundo o número de documentos publicados e proporção teórica da Lei de Lotka clássica. Niterói, RJ, Brasil, 2026

Número de publicações	Número de autores	Proporção observada (%)	Proporção teórica (%)
1	2599	89,2	60,8
2	210	7,2	15,2
3	60	2,1	6,8
4	15	0,5	3,8
5	10	0,3	2,4
6	5	0,2	1,7
7	5	0,2	1,2
8	1	0,0	0,9
9	1	0,0	0,8
10	3	0,1	0,6
11	1	0,0	0,5
12	2	0,1	0,4
13	1	0,0	0,4
18	1	0,0	0,2

Nota: A proporção teórica corresponde ao modelo clássico da Lei de Lotka, calculado por $f(n) = 0,6079/n^2$.

Fonte: dados da pesquisa, 2026.

Geograficamente, os Estados Unidos lideraram a frequência de afiliações (881 ocorrências), seguidos pela China (502 ocorrências). A instituição mais frequente foi a Lanzhou University ($n=102$), seguida pelo National Institute on Aging ($n=47$), University of California ($n=44$), McMaster University ($n=37$) e University of Pennsylvania ($n=35$). Em termos de citação por país de afiliação, os Estados Unidos acumularam 228 citações (média de 9,9 citações por artigo), enquanto a China totalizou 14 citações (média de 2,0 por artigo). Jin Q apresentou o maior número de citações entre os autores, com 195, enquanto Wang Y se destacou pela produtividade.

Estrutura conceitual e mapa temático

O mapa temático bidimensional (Figura 4), construído a partir das palavras-chave dos autores, organizou a estrutura

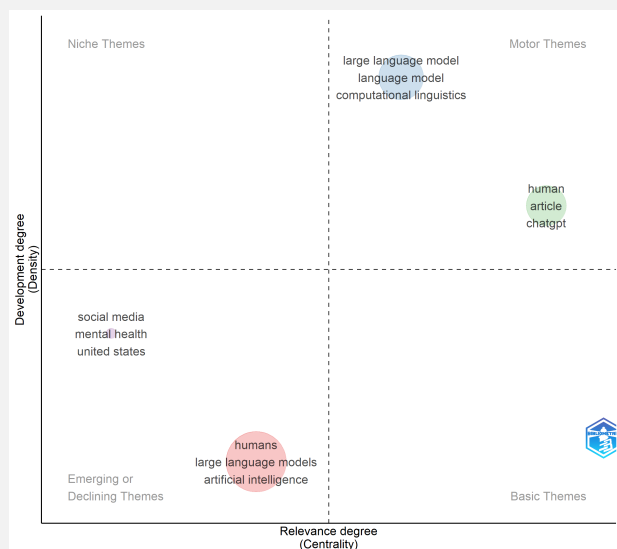
conceitual em quatro agrupamentos, a saber:

Temas Motores (quadrante superior direito): identificaram-se dois clusters bem definidos. O primeiro agrupou os termos “*large language model*”, “*language model*” e “*computational linguistics*”; o segundo reuniu “*ChatGPT*”, “*human*” e “*article*”, indicando o forte desenvolvimento e centralidade metodológica dos LLMs no ecossistema de pesquisa.

Temas Emergentes ou em Declínio (quadrante inferior esquerdo): agruparam-se os termos “*humans*”, “*large language models*” e “*artificial intelligence*”; além do cluster composto por “*social media*”, “*mental health*” e “*United States*”.

Temas de Nicho e Básicos: não foram identificados agrupamentos estruturados nestes quadrantes, reforçando o caráter concentrado e ainda dinâmico da literatura analisada.

Figura 4 – Mapa temático das palavras-chave dos autores associadas à GenAI e aos LLMs como ferramentas de apoio metodológico na pesquisa em saúde. Niterói, RJ, Brasil, 2026



Fonte: dados da pesquisa, gerados no Biblioshiny, 2026.

DISCUSSÃO

Dinâmica de crescimento e estágio de consolidação do campo

Em comparação com bibliometrias mais amplas sobre GenAI na saúde, a produção voltada estritamente ao apoio metodológico à pesquisa ainda é mais delimitada e encontra-se em fase inicial de consolidação. Uma bibliometria sobre ChatGPT na saúde, que incluiu aplicações clínicas, educacionais e de comunicação com pacientes, reuniu 594 publicações e registrou crescimento anual de 91,18% entre 2023 e 2024⁽¹³⁾. Outra análise, voltada ao uso desse mesmo LLM na educação na Ásia, identificou 1.665 documentos publicados entre 2022 e 2025, com taxa anual de crescimento de 545,53%⁽¹⁴⁾. No presente estudo, essa taxa foi de 70,7%, abaixo das observadas nas duas bibliometrias. As diferenças entre bases, períodos e estratégias de busca limitam uma comparação direta entre esses resultados. Ainda assim, os achados mostram que o uso da GenAI e dos LLMs como apoio metodológico à pesquisa em saúde também está em expansão, porém de forma menos acentuada e com maior rigor na seleção de casos de uso.

Informetria das redes de autoria, fontes e cooperação transnacional

A análise das Leis de Bradford e de Lotka, em conjunto com os indicadores de colaboração, aponta na mesma direção. Embora alguns periódicos já concentrem parte relevante das publicações, os documentos estão distribuídos entre 337 fontes, indicando que o tema ainda não possui veículos de disseminação totalmente estabelecidos. A produtividade dos autores também se mostrou fortemente concentrada em contribuições pontuais, com 89,2% deles presentes em apenas um documento, configurando uma

elevada proporção de autores transitórios, e uma queda da produtividade mais acentuada do que a prevista pelo modelo clássico de Lotka ($\beta = 2,87$). A colaboração internacional, presente em 14,7% dos documentos, completa esse quadro ao mostrar que as conexões entre países ainda são pouco frequentes. Em conjunto, esses resultados sugerem que o crescimento da literatura tem ocorrido mais pela entrada de novos autores e pela dispersão do tema entre diferentes periódicos do que pela formação de uma comunidade de pesquisa estável já consolidada.

Hegemonia geográfica e implicações para a equidade científica

A concentração das afiliações nos Estados Unidos e na China deve ser interpretada à luz das diferenças entre países quanto à infraestrutura, à governança, à força de trabalho e à capacidade institucional para a adoção da IA em saúde⁽¹⁵⁾. Embora este estudo não tenha avaliado diretamente essas desigualdades, os resultados mostram que parte relevante do conhecimento sobre o uso da GenAI na pesquisa em saúde está sendo produzida em países com elevada capacidade científica e tecnológica. Esse predomínio pode influenciar as aplicações investigadas, os critérios utilizados para avaliá-las e os problemas considerados prioritários. Tecnologias desenvolvidas e validadas predominantemente nesses contextos podem, portanto, não responder da mesma forma às necessidades de sistemas de saúde e populações que operam em condições de escassez de recursos ou com especificidades epidemiológicas distintas.

Evolução da estrutura e maturidade temática

O mapa temático mostrou que os temas mais desenvolvidos e conectados permanecem centrados na avaliação genérica dos LLMs, com destaque para o ChatGPT. Isso indica que a literatura ainda se organiza em torno das próprias tecnologias e de suas possibilidades de uso, enquanto aplicações metodológicas mais específicas (como a automação de sínteses de evidências ou modelagem de risco de viés) não aparecem ainda como temas consolidados. A ausência de temas básicos ou de nicho e a classificação dos demais agrupamentos como emergentes ou em declínio reforçam esse quadro. Assim, embora a GenAI já esteja sendo incorporada a diferentes etapas da pesquisa em saúde, ainda não se observa uma estrutura temática bem definida em torno de usos metodológicos específicos.

Incertezas metodológicas, integridade e reprodutibilidade

A literatura já descreve o uso da GenAI na busca, triagem, extração de dados, avaliação do risco de viés e síntese de evidências. No entanto, ainda faltam critérios padronizados consistentes para definir em quais etapas da pesquisa os LLMs podem ser utilizados com segurança, quais níveis de erro são aceitáveis e como seu desempenho deve ser auditado e validado^(4,5). Na pesquisa em saúde, essa incerteza é especialmente relevante porque erros na identificação, seleção ou interpretação das evidências podem comprometer revisões que subsidiam protocolos, diretrizes, políticas e decisões de cuidado, com possíveis repercussões

na qualidade e na segurança da assistência⁽⁵⁾.

Por isso, a incorporação da GenAI à pesquisa em saúde não deve ser orientada apenas pelo potencial de agilizar etapas do trabalho, mas também pela capacidade de preservar o rigor, a confiabilidade e a reprodutibilidade das evidências produzidas. O relato de seu uso deve incluir a ferramenta e sua versão, a finalidade de uso, as instruções empregadas, as formas de validação, a verificação dos resultados e os aspectos de privacidade envolvidos⁽¹⁶⁾. Esse detalhamento ajuda a compreender os benefícios e as limitações da tecnologia no contexto em que foi utilizada e permite a auditabilidade das decisões tomadas ao longo do processo. Mesmo com o uso dessas ferramentas, o planejamento, a validação e a interpretação dos resultados continuam sob exclusiva responsabilidade dos pesquisadores⁽¹⁶⁾.

Letramento em IA e transformação da identidade profissional do pesquisador

Por fim, a expansão do uso da GenAI e dos LLMs no processo de pesquisa traz novos desafios para a formação dos pesquisadores. Seu uso requer não apenas o domínio operacional das ferramentas, mas também letramento em IA, capacidade de avaliar criticamente os resultados gerados, reconhecer limitações algorítmicas e justificar decisões metodológicas. Uma revisão sistemática sobre educação e treinamento em IA para trabalhadores da saúde apontou a necessidade de iniciativas formativas que contemplem competências técnicas, avaliação crítica, aspectos éticos e uso seguro dessas tecnologias⁽¹⁷⁾, enquanto outro estudo mostrou que a identidade profissional pode estar relacionada à prontidão para o uso da IA entre trabalhadores da saúde⁽¹⁸⁾. Esses achados sugerem que a formação dos pesquisadores precisa acompanhar essas mudanças, considerando não apenas as competências necessárias para o uso da GenAI, mas também a compreensão de seu papel e de sua responsabilidade diante das novas formas de produzir conhecimento^(17,18).

Limitações do estudo

Como limitações, os resultados refletem as bases consultadas, a estratégia de busca e o recorte temporal adotado. Os indicadores de citação foram obtidos dos metadados importados e devem ser interpretados como medidas descritivas do corpus, pois a cobertura varia entre as fontes. A utilização de um LLM como apoio à triagem também deve ser considerada, uma vez que classificações automatizadas estão sujeitas a erros e podem ter influenciado a composição do corpus, apesar da supervisão humana. Ao mesmo tempo, seu uso no próprio processo de seleção permitiu explorar, na prática, uma das aplicações metodológicas que constituem o objeto deste estudo. Além disso, como a produção sobre o tema cresce rapidamente, o panorama apresentado está condicionado ao período analisado e pode se modificar com a publicação de novos estudos.

CONCLUSÃO

Este estudo mostrou que o uso da GenAI e dos LLMs

como apoio metodológico à pesquisa em saúde constitui um campo recente e em expansão, mas ainda marcado pela dispersão das publicações entre autores e periódicos, baixa colaboração internacional e ausência de uma estrutura temática consolidada em torno de aplicações metodológicas específicas. Os achados indicam que a incorporação dessas tecnologias avança mais rapidamente do que a organização de uma base científica e regulatória capaz de orientar seus usos na pesquisa em saúde.

Recomenda-se que estudos futuros validem o desempenho dessas ferramentas em etapas metodológicas específicas, que seu uso seja relatado de forma padronizada e transparente, com supervisão humana, e que o letramento em IA seja incorporado à formação de pesquisadores, inclusive na enfermagem. Também se recomendam o fortalecimento de redes colaborativas, a ampliação dos estudos em países do Sul Global, que tendem a incorporar tecnologias de IA desenvolvidas em outros contextos, e a atualização periódica deste mapeamento. Essas medidas são necessárias porque erros na busca, seleção, análise e síntese de evidências podem repercutir na qualidade e na segurança do cuidado, bem como nas diretrizes e políticas de saúde.

CONFLITO DE INTERESSES

Os autores declaram não haver conflito de interesses.

FINANCIAMENTO

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL

Ferramentas de GenAI foram utilizadas em diferentes etapas deste estudo. No desenvolvimento do fluxo computacional em Python, foram utilizados o Codex e o Claude Code. A triagem de títulos e resumos foi realizada com apoio do Claude Opus 5, a partir de critérios de elegibilidade e de um prompt definidos e refinados pela pesquisadora. O ChatGPT e o Claude também foram utilizados como apoio à organização, redação e revisão linguística do manuscrito.

O uso dessas ferramentas ocorreu sob supervisão humana integral. As classificações geradas pelo Claude durante a triagem foram submetidas à supervisão da pesquisadora, com avaliação prioritária dos registros classificados como incertos e ajuste das decisões quando necessário. A definição dos critérios de elegibilidade, as decisões metodológicas, a interpretação dos resultados e as conclusões permaneceram sob responsabilidade dos pesquisadores, que revisaram criticamente a versão final e assumem integral responsabilidade pelo conteúdo apresentado.

REFERÊNCIAS

1. Truhn D, Reis-Filho JS, Kather JN. Large language models should be used as scientific reasoning engines, not knowledge databases. *Nat Med*. 2023;29(12):2983-2984. <https://doi.org/10.1038/s41591-023-02594-z>. PMID: 37853138.
2. Wei C, Allot A, Lai P, Leaman R, Tian S, Luo L, et al. PubTator 3.0: an AI-powered literature resource for unlocking biomedical knowledge. *Nucleic Acids Research*. 2024;52(W1):W540-W546. <https://doi.org/10.1093/nar/gkae235>. PMID: 38572754.
3. Jin Q, Leaman R, Lu Z. PubMed and beyond: biomedical literature search in the age of artificial intelligence. *eBioMedicine*. 2024;100:104988. <https://doi.org/10.1016/j.ebiom.2024.104988>. PMID: 38306900.
4. Lieberum J, Toews M, Metzendorf M, Heilmeyer F, Siemens W, Haverkamp C, et al. Large language models for conducting systematic reviews: on the rise, but not yet ready for use—a scoping review. *Journal of Clinical Epidemiology*. 2025;181:111746. <https://doi.org/10.1016/j.jclinepi.2025.111746>. PMID: 40021099.
5. Siemens W, von Elm E, Binder H, Böhringer D, Eisele-Metzger A, Gartlehner G, et al. Opportunities, challenges and risks of using artificial intelligence for evidence synthesis. *BMJ EBM*. 2025;30(6):381-384. <https://doi.org/10.1136/bmjebm-2024-113320>. PMID: 39788693.
6. Aria M, Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*. 2017;11(4):959-975. <https://doi.org/10.1016/j.joi.2017.08.007>.
7. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*. 2021;133:285-296. <https://doi.org/10.1016/j.jbusres.2021.04.070>.
8. Bradford SC. Sources of information on specific subjects. *Engineering*. 1934;137:85-86.
9. Lotka AJ. The frequency distribution of scientific productivity. *J Wash Acad Sci*. [Internet]. 1926 [citado 2026 set. 12];16(12):317-23. Disponível em: <https://www.jstor.org/stable/24529203>.
10. Zipf GK. Human behavior and the principle of least effort: an introduction to human ecology. Cambridge: Addison-Wesley; 1949.
11. Gerasimov I, KC B, Mehrabian A, Acker J, McGuire MP. Comparison of datasets citation coverage in Google Scholar, Web of Science, Scopus, Crossref, and DataCite. *Scientometrics*. 2024;129(7):3681-3704. <https://doi.org/10.1007/s11192-024-05073-5>.
12. Brasil. Conselho Nacional de Saúde. Resolução nº 510, de 7 de abril de 2016. Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais cujos procedimentos metodológicos envolvam a utilização de dados diretamente obtidos com os participantes ou de informações identificáveis ou que possam acarretar riscos maiores do que os existentes na vida cotidiana, na forma definida nesta Resolução [Internet]. Brasília: Ministério da Saúde; 2016 [citado 2026 set. 12]. Disponível em: <https://www.gov.br/conselho-nacional-de-saude/pt-br/atos-normativos/resolucoes/2016/resolucao-no-510.pdf/view>.
13. Amir AS. Revolutionizing healthcare with ChatGPT: current advancements, challenges, and future directions. *Discov Artif Intell*. 2026;6(1):582. <https://doi.org/10.1007/s44163-026-01506-7>.
14. Irwanto I. Research on ChatGPT in education in Asia: a bibliometric analysis (2022–2025). *Discov Artif Intell*. 2025;5(1):244. <https://doi.org/10.1007/s44163-025-00505-4>.
15. World Health Organization, Regional Office for Europe. Artificial intelligence is reshaping health systems: state of readiness across the European Union [Internet]. Copenhagen: WHO Regional Office for Europe; 2026 [citado 2026 set. 12]. Disponível em: <https://www.who.int/europe/publications/iit-em/WHO-EURO-2026-12707-52481-81471>
16. Luo X, Tham YC, Giuffrè M, Ranisch R, Daher M, Lam K, et al. Reporting guideline for the use of Generative Artificial intelligence tools in MEDical Research: the GAMER Statement. *BMJ EBM*. 2025;30(6):390-400. <https://doi.org/10.1136/bmjebm-2025-113825>. PMID: 40360239.
17. Woods L, Lyons K, Van Der Vegt A, Olsen Q, Huang W, Khor JS, et al. Assessing the effectiveness of artificial intelligence education and training for healthcare workers: a systematic review. *BMC Med Educ*. 2026;26(1). <https://doi.org/10.1186/s12909-026-08969-3>. PMID: 41807961.
18. Ta'an W, Damrah S, Al-Hammouri MM, Williams B. Professional identity and its relationships with AI readiness and inter-professional collaboration. *PLoS One*. 2025;20(5):e0322794. <https://doi.org/10.1371/journal.pone.0322794>. PMID: 40378129.

CONTRIBUIÇÃO DE AUTORIA

Concepção do estudo: Almeida YS, Cortez EA, Valente GSC.

Obtenção de dados: Almeida YS.

Análise de dados: Almeida YS.

Interpretação dos dados: Almeida YS.

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.



Copyright © 2026 Online Brazilian Journal of Nursing

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.