Rev Obstet Ginecol Venez. 2026; 86 (2): 322-334.
https://doi.org/10.51288/00860214

Herramientas de inteligencia artificial para la investigación científica biomédica: aplicaciones prácticas en búsqueda, redacción y revisión de manuscritos

Diana Ivy Anca Vega.[1]

RESUMEN

La integración de la inteligencia artificial generativa en el trabajo académico biomédico plantea oportunidades y riesgos que merecen un análisis cuidadoso. Los grandes modelos de lenguaje y las herramientas de búsqueda semántica, mapeo de literatura, redacción asistida y edición lingüística pueden apoyar tareas frecuentes del investigador, como formular preguntas, localizar referencias, resumir evidencia, organizar manuscritos, mejorar la claridad y revisar documentos antes del envío editorial. Su utilidad depende, en última instancia, de la supervisión crítica del autor. Los modelos generativos pueden producir información plausible pero incorrecta, incluidas referencias inexistentes o inexactas, por lo que no deben reemplazar la lectura de las fuentes originales, la verificación bibliográfica ni la evaluación metodológica de la evidencia. Esta revisión narrativa describe herramientas aplicables a distintas etapas del trabajo científico biomédico, con énfasis en la utilidad práctica, las limitaciones, la trazabilidad de las fuentes y la declaración transparente de su uso.

***Palabras clave**: Inteligencia artificial, Procesamiento de lenguaje natural, Investigación biomédica, Escritura científica, Recuperación de información, Publicaciones periódicas.*

Artificial intelligence tools for biomedical scientific research: practical applications in manuscript search, writing and review

SUMMARY

The integration of generative artificial intelligence into biomedical academic work raises opportunities and risks that warrant careful consideration. Large language models, semantic search engines, literature mapping systems, writing assistants, and language editing tools may support common research tasks, including question formulation, reference retrieval, evidence summarization, manuscript organization, language refinement, and pre-submission review. Their usefulness ultimately rests on critical human oversight. Generative models may produce plausible but incorrect information, including fabricated or inaccurate references; they should not replace reading original sources, verifying bibliographic data, or critically appraising the evidence. This narrative review describes artificial intelligence tools applicable to different stages of biomedical research, with emphasis on practical usefulness, limitations, source traceability, and transparent disclosure of their use in scientific publications.

***Keywords:** Artificial intelligence, Natural language processing, Biomedical research, Scientific writing, Information storage and retrieval, Periodicals as topic.*

INTRODUCCIÓN

La inteligencia artificial (IA) ha pasado de ser un recurso asociado con laboratorios especializados o con el análisis masivo de datos a integrarse en tareas cotidianas de lectura, escritura y organización del conocimiento. En términos generales, puede entenderse como el conjunto de métodos computacionales capaces de realizar tareas que habitualmente requieren funciones cognitivas humanas, como reconocer patrones, procesar lenguaje, aprender de datos o apoyar la toma de decisiones. En investigación biomédica, este cambio no se limita al desarrollo de algoritmos diagnósticos o de modelos predictivos, pues también alcanza actividades decisivas para la producción científica, como buscar literatura, resumir artículos, ordenar argumentos, redactar borradores, revisar el estilo, traducir textos y preparar manuscritos para publicación (1, 2).

Dentro de este campo, la inteligencia artificial generativa ocupa un lugar particular. Estos sistemas

[1]Médico especialista en Ginecología y Obstetricia. Máster en Inteligencia Artificial Aplicada a la Sanidad. Policlínica La Arboleda / Grupo Médico Santa Paula. Caracas, Venezuela.
Correo para correspondencia: dianaanca@hotmail.com

Forma de citar este artículo: Anca Vega DI. Herramientas de inteligencia artificial para la investigación científica biomédica: aplicaciones prácticas en búsqueda, redacción y revisión de manuscritos. Rev Obstet Ginecol Venez. 86(2):322-334. DOI: 10.51288/00860214

producen contenido nuevo (texto, imágenes, audio, código o respuestas conversacionales) a partir de patrones aprendidos en grandes volúmenes de datos (1, 3). En medicina, salud pública e investigación médica se han explorado aplicaciones en síntesis de información, traducción, adaptación de mensajes para distintas audiencias, extracción de datos y apoyo preliminar a la comunicación científica, siempre bajo responsabilidad humana (1, 3).

El interés por estas herramientas se explica, en parte, por las exigencias actuales de la escritura científica. El investigador biomédico debe combinar dominio temático, criterio metodológico, capacidad de síntesis, precisión bibliográfica y claridad lingüística en un entorno de crecimiento acelerado de la literatura. PubMed contiene decenas de millones de registros y suma más de un millón de referencias cada año, de modo que una búsqueda simple puede recuperar cientos o miles de resultados, de los cuales solo una fracción llega a revisarse con detalle (4). Que esta incorporación ya alcanza una escala medible lo confirma un análisis de más de 15 millones de resúmenes biomédicos de PubMed publicados entre 2010 y 2024: se estimó que al menos el 13,5 % de los resúmenes de 2024 fueron procesados con grandes modelos de lenguaje (LLM), con proporciones que llegaron al 40 % en algunos subcorpus, países y revistas, y con un efecto medible sobre el vocabulario científico biomédico (5). En ese contexto, la aceleración sostenida en el desarrollo de los modelos, cada generación con mayor capacidad de contexto, razonamiento y acceso a fuentes externas, convierte la evaluación crítica de sus posibilidades y sus límites en una competencia cada vez más necesaria. La IA puede servir de apoyo para explorar, filtrar y organizar información, pero no sustituye a las fuentes originales.

La evidencia disponible aconseja prudencia. La IA puede acelerar tareas y mejorar la claridad de los textos, aunque también puede introducir errores, sesgos, referencias inexistentes o interpretaciones no justificadas cuando se usa sin supervisión (1, 3, 6). El problema de las citas resulta especialmente sensible. Chelli y cols. (6) evaluaron GPT-3.5, GPT-4 y Bard en la generación de referencias para revisiones sistemáticas sobre patología del manguito rotador, y documentaron tasas de alucinación del 39,6 %, 28,6 % y 91,4 %, respectivamente, con baja precisión en los tres modelos. Aunque el campo evaluado fue acotado, estos hallazgos refuerzan una regla editorial básica: los modelos generativos no deben usarse como verificadores bibliográficos finales ni como única vía para construir una lista de referencias, con independencia de las mejoras que aporten sus versiones sucesivas.

La discusión editorial avanza en la misma dirección. La declaración GAMER (Generative Artificial intelligence tools in MEdical Research) propone nueve ítems para informar el uso de IA generativa en investigación médica, entre ellos la identificación de la herramienta y su versión, el rol que tuvo en el estudio, las secciones asistidas, la forma de verificación del contenido, la protección de datos y el posible impacto sobre la interpretación o las conclusiones (3). Esta orientación traslada el debate desde la simple aceptación o rechazo de las herramientas hacia una pregunta más útil para autores y revistas: cómo informar su uso de manera transparente y trazable.

Para los investigadores biomédicos hispanohablantes, y en particular para quienes publican desde América Latina, el tema adquiere una dimensión adicional. Varias barreras de publicación se relacionan con el acceso a literatura actualizada, el dominio del inglés académico, la disponibilidad de apoyo metodológico y la familiaridad con las normas editoriales internacionales. La IA puede ayudar a acortar parte de ese recorrido, siempre que el autor conserve la responsabilidad de formular buenas preguntas, contrastar fuentes y asumir el contenido final del texto.

Esta revisión describe, con orientación práctica, herramientas y recursos de IA que pueden apoyar distintas etapas del trabajo científico biomédico: la formulación de ideas, la búsqueda y organización de literatura, la selección y síntesis de evidencia, la gestión bibliográfica, la redacción, la edición, la revisión final del manuscrito y la declaración transparente del uso de IA. El propósito es ofrecer al médico y al investigador elementos para decidir qué herramientas pueden ser útiles, cuándo emplearlas, qué esperar de ellas y qué límites conviene no perder de vista.

MÉTODOS. ESTRATEGIA DE BÚSQUEDA PARA LA REVISIÓN NARRATIVA

Se elaboró una revisión narrativa orientada a la utilidad práctica de la IA en la investigación biomédica. Este diseño se consideró adecuado porque el tema evoluciona con rapidez, combina literatura biomédica, informática, comunicación científica y políticas editoriales, y no busca responder una pregunta clínica de eficacia mediante la estimación cuantitativa de resultados. En revisiones sobre IA generativa y comunicación científica se ha señalado que el formato narrativo permite integrar evidencia reciente, documentos de orientación y experiencias de aplicación cuando el campo está en desarrollo y la evidencia es heterogénea (1).

La búsqueda se realizó en PubMed/MEDLINE y se complementó con motores de búsqueda semántica y con la revisión manual de las referencias de los artículos seleccionados. Se consideraron documentos publicados entre enero de 2023 y mayo de 2026, en inglés y en español, mediante combinaciones de términos como "*artificial intelligence*", "*generative artificial intelligence*", "*large language models*", "*scientific writing*", "*biomedical literature search*", "*systematic review*" y sus equivalentes en español. Se identificaron cinco grupos de documentos: artículos sobre IA generativa y grandes modelos de lenguaje en medicina; revisiones o análisis sobre búsqueda biomédica asistida por IA; estudios sobre desempeño, errores y alucinaciones de modelos generativos; documentos editoriales o guías de reporte sobre la declaración del uso de IA; y trabajos sobre escritura científica, edición lingüística y comunicación académica.

Se priorizaron fuentes biomédicas revisadas por pares, estudios metodológicos, revisiones, guías de reporte y documentos de organizaciones o revistas reconocidas. Las fuentes comerciales o las páginas institucionales de las herramientas se utilizaron solo para describir funcionalidades declaradas por los proveedores, no como evidencia independiente de desempeño o de superioridad frente a otras alternativas. No se aplicó la metodología PRISMA, ya que el objetivo no fue realizar una revisión sistemática ni estimar el efecto de una intervención. Como criterio editorial, cada recomendación operativa se vinculó con una fuente verificable cuando estuvo disponible; cuando una herramienta se menciona por su funcionalidad y sin evaluación académica suficiente, se presenta como recurso exploratorio y no como intervención validada.

RESULTADOS DE LA REVISIÓN NARRATIVA

Herramientas de IA para formular ideas, delimitar preguntas y planificar el manuscrito

La primera utilidad práctica de la IA para el investigador no reside necesariamente en escribir el artículo, sino en ordenar el pensamiento previo a la escritura. Un modelo conversacional puede ayudar a transformar una inquietud amplia en una pregunta más precisa, sugerir términos alternativos, identificar posibles enfoques de revisión, proponer un esquema inicial o señalar vacíos argumentales en una idea todavía inmadura (1, 7). Esta función resulta especialmente valiosa cuando el

investigador cuenta con experiencia clínica pero aún no la ha traducido en una pregunta publicable.

Herramientas como ChatGPT de OpenAI, Claude de Anthropic, Gemini de Google, Microsoft Copilot o DeepSeek pueden emplearse en esta etapa como asistentes de diálogo. Su valor no está en que descubran por sí mismas una pregunta científica, sino en que permiten ensayar formulaciones, comparar versiones, anticipar cómo podría organizarse un manuscrito y convertir una pregunta amplia en variantes estructuradas según PICO (Paciente/Población, Intervención, Comparación, Resultado), PECO (Población, Exposición, Comparación, Resultado) o SPIDER Muestra, Fenómeno de Interés, Diseño, Evaluación, Tipo de investigación). La revisión de Yoga Ratnam (1) ofrece un ejemplo de integración razonable: la IA generativa se empleó como apoyo para organizar temas y mejorar la redacción, pero los resultados se revisaron, corrigieron y validaron por el autor antes de incorporarse al trabajo final.

Este uso temprano exige cautela en dos sentidos. En primer lugar, la pregunta científica debe responder a una necesidad real, ser pertinente para la audiencia de la revista y tener respaldo suficiente en la literatura. En segundo lugar, estos modelos tienden a ser complacientes con los planteamientos del investigador y a confirmar las hipótesis propuestas en lugar de cuestionarlas, de modo que el espíritu crítico debe prevalecer sobre la fluidez de las respuestas. Los grandes modelos de lenguaje producen respuestas probables a partir de patrones de texto, no evaluaciones metodológicas independientes en sentido estricto (2, 3). Una forma práctica de trabajo consiste en pedir al modelo tres productos: una lista de posibles preguntas de investigación; una tabla que relacione cada pregunta con el tipo de artículo, la población, el desenlace y las fuentes necesarias; y un esquema de manuscrito con subtítulos tentativos. Ese material no debe copiarse directamente al artículo, aunque sí puede servir como borrador de trabajo. El *prompting*, entendido como el diseño preciso de las instrucciones enviadas a un modelo de lenguaje, constituye la base de cualquier interacción efectiva. La calidad de la respuesta depende directamente de la calidad de la instrucción: un prompting eficaz debe proporcionar un contexto claro, definir el rol del modelo, delimitar el alcance de la tarea y establecer restricciones específicas. En el ámbito académico, un prompting bien estructurado es fundamental para evitar respuestas genéricas o alucinaciones, asegurando que el modelo se mantenga alineado con los objetivos de investigación y el rigor científico requerido.

Búsqueda bibliográfica: de PubMed a los motores semánticos

La búsqueda de literatura sigue siendo una etapa crítica del trabajo científico. Un manuscrito puede estar bien redactado y, aun así, debilitarse si su fundamento bibliográfico es incompleto, desactualizado o poco pertinente. En investigación biomédica, PubMed conserva un papel central por su cobertura, su trazabilidad y su familiaridad para médicos e investigadores; como ya se señaló, el volumen actual de la base hace inviable el cribado manual exhaustivo en la práctica cotidiana (4).

Las herramientas de IA no sustituyen a PubMed, Embase, Scopus, Web of Science o la Cochrane Library cuando se requiere una búsqueda formal y reproducible. Su papel más razonable es complementario, pues ayudan a explorar conceptos, identificar términos, resumir resultados iniciales, ubicar artículos relacionados y reconocer líneas de investigación que una búsqueda por palabras clave podría pasar por alto (4, 8). Una cosa es explorar la literatura para orientar una revisión narrativa y otra distinta es construir una estrategia sistemática, documentada y reproducible. En un editorial reciente, Sen (8) repasó la evolución de las herramientas de búsqueda semántica y minería de literatura entre 2020 y 2025, y subrayó que su utilidad para reducir el tiempo de exploración inicial no elimina la necesidad de validar cada fuente seleccionada.

Entre las herramientas de búsqueda semántica y de respuesta basada en literatura se encuentran Semantic Scholar, Elicit, Consensus, Perplexity y OpenRead. Elicit ha sido evaluada como apoyo en revisiones sistemáticas, lo que refleja el interés creciente por usar IA en las etapas iniciales de identificación y organización de la evidencia (9). Plataformas como OpenEvidence y Vera Health se presentan como sistemas de consulta clínica basada en fuentes; su utilidad está en orientar una primera aproximación a una pregunta clínica, si bien, cuando la evidencia disponible procede sobre todo de documentación institucional o comercial, conviene tratarlas como recursos exploratorios y no como fuentes primarias (10, 11).

ResearchRabbit, Connected Papers, Litmaps e Iris.ai permiten partir de una o varias referencias clave para explorar redes de artículos relacionados. Son útiles para identificar autores, líneas de trabajo y publicaciones previas o posteriores a un estudio relevante, aunque pueden favorecer la literatura más citada, en inglés o procedente de circuitos académicos de mayor visibilidad. Scite.ai aporta una función distinta, ya que permite revisar el contexto en que un artículo ha sido citado por otros trabajos y ayuda a distinguir si una referencia se usó como apoyo, contraste o simple mención. Aun así, ese análisis no reemplaza la lectura crítica del artículo original.

En términos prácticos, una búsqueda asistida por IA puede organizarse en tres pasos. Primero, definir la pregunta y realizar una búsqueda inicial en PubMed con términos MeSH y sinónimos. Segundo, usar herramientas semánticas para explorar relaciones, artículos relacionados y respuestas preliminares. Tercero, verificar de forma manual cada fuente seleccionada en PubMed, Crossref, el sitio de la revista o bases académicas confiables. La IA puede acelerar la exploración, pero la trazabilidad bibliográfica sigue dependiendo del investigador.

Selección, cribado y síntesis de evidencia

Tras la búsqueda aparece una etapa menos visible, pero igual de importante: decidir qué literatura se usará y por qué, con base en la relevancia, la actualidad, la calidad metodológica y la utilidad para el argumento. En las revisiones sistemáticas, el proceso exige criterios explícitos de inclusión y exclusión, revisión por pares de los registros, extracción de datos y evaluación formal del riesgo de sesgo. En ambos tipos de revisión, la IA puede apoyar tareas repetitivas, pero no debe reemplazar la decisión metodológica.

Rayyan, Covidence, EPPI-Reviewer, DistillerSR y ASReview se utilizan para organizar el cribado de títulos y resúmenes, detectar duplicados, etiquetar registros y priorizar referencias. Con los grandes modelos de lenguaje ha crecido el interés por usar sistemas como ChatGPT o GPT-4 en las etapas de búsqueda, cribado, extracción y evaluación preliminar de la evidencia (12 - 15). Estos recursos, no obstante, deben entenderse como apoyo operativo en un campo de evolución rápida, donde las características, las fuentes de información, las políticas de acceso y el rendimiento de las herramientas pueden cambiar en poco tiempo. La decisión de inclusión, la extracción final y la evaluación del riesgo de sesgo requieren revisión humana.

Gwon y cols. (13) evaluaron ChatGPT y Microsoft Bing AI para búsquedas de literatura científica en revisiones sistemáticas, lo que ilustra el interés por determinar hasta qué punto estos modelos pueden apoyar tareas tradicionalmente realizadas por personas. Los estudios sobre extracción de datos y cribado automatizado sugieren una utilidad potencial, aunque sus resultados deben interpretarse con cautela y verificarse contra el texto completo de los artículos (14, 15). La extracción de datos exige un cuidado especial, porque los modelos pueden omitir datos, confundir denominadores, mezclar información entre estudios o transformar conclusiones cautelosas en afirmaciones

más fuertes. Para la evaluación del riesgo de sesgo, los modelos pueden ayudar a localizar información relevante o a recordar los dominios de herramientas como ROBINS-I, pero no deben emitir la clasificación final sin revisión metodológica experta (16).

Gestión bibliográfica y verificación de referencias

Zotero, Mendeley y EndNote facilitan almacenar artículos, organizar colecciones, insertar citas y aplicar estilos como Vancouver. Un gestor bibliográfico, sin embargo, no verifica por sí mismo que una referencia exista, que el DOI sea correcto o que el artículo respalde la afirmación citada. Esa verificación sigue siendo una tarea editorial del autor. Antes del envío, cada referencia debe revisarse en PubMed, Crossref, DOI.org o el sitio de la revista, y deben confirmarse el título, los autores, el año, la revista, el volumen, las páginas, el DOI o el PMID, y la correspondencia entre la cita y la afirmación. Esta precaución cobra particular importancia tras el uso de modelos generativos, dado el riesgo documentado de referencias alucinadas o inexactas (6).

La auditoría bibliográfica debe responder dos preguntas distintas: ¿la referencia existe? y ¿la referencia corresponde a esta afirmación? La segunda suele ser la más descuidada y, probablemente, la más importante. Una verificación práctica puede realizarse en tres niveles. El nivel bibliográfico confirma que la referencia existe y que sus datos son correctos. El nivel semántico comprueba que la afirmación del manuscrito corresponde al contenido del artículo citado. El nivel metodológico valora si el diseño del estudio permite sostener la fuerza de la afirmación. Un comentario editorial puede respaldar una recomendación de política, pero no demostrar la eficacia de una herramienta; una página comercial puede describir funcionalidades, aunque no validar su rendimiento; y un estudio de prueba de concepto puede mostrar factibilidad sin que ello permita generalizar los resultados.

Herramientas para redacción, edición y traducción del manuscrito

La IA generativa puede apoyar la escritura, aunque conviene precisar de qué manera: sirve para ordenar, reformular y revisar, no para producir por sí sola el contenido científico final. ChatGPT, Claude, Gemini, Copilot y herramientas similares resultan útiles para convertir notas en un primer esquema, mejorar la transición entre párrafos, reducir repeticiones o comprobar si una sección responde al objetivo planteado (17). En los artículos biomédicos, su mejor uso es editorial, es decir, mejorar la claridad sin inventar evidencia. Una instrucción acotada, como mejorar la claridad sin cambiar el sentido, reducir redundancias o identificar afirmaciones que requieren cita, conserva el control del autor. Las instrucciones abiertas, como pedir al modelo que redacte la discusión completa a partir de unas referencias, resultan más arriesgadas, porque el modelo puede mezclar hallazgos, exagerar conclusiones o atribuir resultados a artículos que no los contienen (6, 17).

Junto a los modelos conversacionales de uso general, existe una categoría emergente de plataformas integradas, diseñadas específicamente para la escritura académica, que reúnen en un solo entorno funciones de redacción asistida, gestión de citas, revisión lingüística y comprobación previa al envío editorial. Estas herramientas pueden resultar útiles para investigadores que buscan un flujo de trabajo unificado, si bien su evaluación académica independiente es aún limitada en comparación con su documentación comercial. Deben considerarse recursos exploratorios, útiles para tareas operativas concretas, pero sin evidencia suficiente para recomendarlas sobre alternativas establecidas ni para sustituir la verificación bibliográfica, la lectura crítica de las fuentes o la responsabilidad del autor sobre el contenido final.

Las herramientas de edición lingüística cumplen un papel más específico. Grammarly, Writefull,

Paperpal, DeepL Write, LanguageTool y Microsoft Editor pueden mejorar la gramática, la puntuación, la fluidez y la consistencia del estilo. Para los autores hispanohablantes que publican en inglés, DeepL y los modelos conversacionales pueden servir como apoyo de traducción inicial; esta utilidad se ha señalado para investigadores no anglófonos, en quienes la IA puede reducir las barreras idiomáticas de la escritura científica (18 - 20). El límite es claro: una traducción correcta en apariencia puede alterar matices técnicos, modificar el grado de certeza o convertir una asociación en una afirmación causal. La revisión por un autor con dominio del contenido médico sigue siendo obligatoria.

Chetwynd (21) advierte sobre la necesidad de un uso ético y crítico de la IA en la escritura científica, sobre todo para evitar que la fluidez formal oculte debilidades de razonamiento, de atribución o de responsabilidad autoral. En una revisión narrativa esto resulta especialmente relevante, pues el lector no necesita párrafos impecables pero vacíos, sino una síntesis con criterio, jerarquía y lectura crítica de la evidencia.

Revisión final del manuscrito

La revisión final es una etapa en la que la IA puede ahorrar tiempo si se emplea como lista de comprobación. Un modelo generativo puede detectar frases repetidas, cambios de terminología, saltos lógicos, objetivos que no coinciden con la conclusión o afirmaciones que aparecen sin referencia. También puede comparar el manuscrito con las instrucciones para autores, revisar la extensión del resumen, sugerir palabras clave o identificar secciones faltantes. Leung y cols. (17) señalan que las herramientas de IA pueden participar en tareas editoriales como la revisión técnica, la validación de referencias, los resúmenes, las traducciones y el apoyo al procesamiento editorial, siempre con supervisión humana.

Una estrategia útil consiste en realizar tres revisiones separadas. La primera se centra en la coherencia entre el objetivo, el desarrollo y la conclusión. La segunda se ocupa de la evidencia y verifica que cada afirmación importante tenga una fuente que realmente la respalde. La tercera atiende a la forma y revisa el estilo Vancouver, las tablas, las abreviaturas, las palabras clave, el resumen y las normas de la revista. La revisión de las referencias merece un trato independiente, pues antes de enviar el manuscrito cada cita debe confirmarse en PubMed, Crossref o el sitio de la revista. Conviene además evitar la introducción de manuscritos confidenciales, datos identificables de pacientes o documentos de terceros en herramientas públicas sin autorización. La guía de JAMA para autores, revisores y editores exige transparencia cuando se utilizan IA, modelos de lenguaje o chatbots para crear o editar contenido, y recuerda que los autores deben asumir la responsabilidad por la integridad del material generado con estas herramientas (22).

Declaración de uso de inteligencia artificial

La declaración de uso de IA debe ser breve, concreta y verificable. Siguiendo los nueve elementos propuestos por GAMER (3), una declaración suficiente para un artículo de revisión como este puede indicar la herramienta utilizada y su versión, la función que cumplió (organización temática, revisión de claridad o edición preliminar), que las sugerencias generadas se revisaron y contrastaron con las fuentes originales, que no se introdujeron datos identificables y que el autor asume la responsabilidad plena por el contenido final.

Esta declaración no exime de revisar las normas de cada revista. La heterogeneidad de las políticas editoriales es notable. Ganjavi y cols. (23) analizaron las instrucciones de los 100 mayores editores académicos y de las 100 revistas más citadas, y encontraron que el 24 % de los editores y el 87 % de las revistas de mayor rango habían publicado directrices sobre el uso de IA, con una variabilidad considerable en el nivel de detalle exigido. Análisis más recientes confirman esa

dispersión entre las revistas médicas (24, 25). Por ello, además de aplicar una guía general como GAMER, los autores deben revisar las instrucciones actualizadas de la revista de destino. El ICMJE (International Committee of Medical Journal Editors) establece que las tecnologías asistidas por IA no deben figurar como autoras y que los autores humanos son responsables de revisar, editar y garantizar la exactitud del contenido generado con estas herramientas (26).

Riesgos y límites prácticos

Los riesgos del uso de IA en investigación biomédica pueden agruparse en cuatro categorías que conviene distinguir, porque exigen estrategias de control diferentes. El riesgo de contenido se refiere a respuestas plausibles pero falsas, como referencias inexistentes, datos bibliográficos incorrectos o resúmenes que omiten matices relevantes (6). El riesgo metodológico abarca los sesgos derivados de los corpus de entrenamiento, el predominio de la literatura en inglés y la sobreinterpretación de la evidencia débil (2, 27). El riesgo ético cubre la exposición de datos clínicos identificables, manuscritos no publicados, revisiones por pares o documentos institucionales reservados en plataformas externas (3, 22). Por último, el riesgo editorial incluye la dependencia intelectual, los textos formalmente correctos pero conceptualmente vacíos, la declaración insuficiente del uso de IA y la pérdida de trazabilidad entre la afirmación y la fuente; este riesgo abarca incluso la facilitación de prácticas editoriales fraudulentas, como la escritura fantasma o la participación en circuitos de paper mills, cuando la herramienta se usa sin trazabilidad ni supervisión (21, 28). Ramoni y cols. (28) resumen esta tensión al diferenciar los usos legítimos, como la mejora de la claridad, la traducción, la exploración de literatura y la revisión de estilo, de los usos engañosos, como la generación de datos falsos, la redacción completa sin supervisión y el fraude editorial. El debate, por tanto, no se reduce a si la IA se usa, sino a cómo se usa y con qué grado de supervisión.

Por otra parte, las críticas y objeciones vertidas sobre el desempeño de estos modelos han sido atendidas por los desarrolladores en las actualizaciones de sus versiones, logrando una mayor eficiencia y utilidad en las labores encomendadas. No obstante, es imperativo señalar que, hasta la fecha, no existen estudios académicos independientes que respalden de forma rigurosa esta evolución; la percepción de mejora se sustenta principalmente en benchmarks y métricas de desempeño no validadas externamente, que, si bien ofrecen una orientación sobre el progreso tecnológico, carecen aún del escrutinio científico necesario para confirmar su superioridad real en el contexto de la investigación biomédica.

Propuesta de flujo práctico para el investigador biomédico

Un flujo razonable puede organizarse en cinco momentos. Primero, definir la pregunta con criterio clínico o metodológico propio y usar la IA solo para explorar formulaciones alternativas. Segundo, iniciar la búsqueda en PubMed u otra base biomédica y ampliar la exploración con herramientas semánticas como Elicit, Consensus, Semantic Scholar, OpenEvidence o Vera Health, tratando estas últimas como recursos exploratorios cuando no exista validación independiente suficiente. Tercero, organizar las referencias en Zotero, Mendeley o EndNote y verificar cada fuente antes de citarla. Cuarto, usar la IA para mejorar la estructura, la claridad y el estilo, no para reemplazar la interpretación de los artículos. Quinto, antes del envío, revisar la coherencia, las citas, las normas de la revista y la declaración de uso de IA.

La herramienta funciona mejor cuando el investigador sabe qué quiere preguntar, qué fuente debe consultar y qué nivel de evidencia necesita. Para elegir entre herramientas, los criterios mínimos son la trazabilidad de las fuentes, la posibilidad de verificar las respuestas, la protección de datos, la facilidad de uso y la utilidad real para la tarea que se desea resolver (3, 17, 22). Una

herramienta puede ser excelente para resumir artículos y, a la vez, deficiente para construir una estrategia de búsqueda; otra puede mejorar la gramática sin aportar nada a la calidad metodológica del manuscrito. El primer criterio sigue siendo la verificabilidad, de modo que las herramientas que ofrecen enlaces a artículos, DOI o contexto de citación comprobables tienen ventaja sobre los sistemas que entregan respuestas cerradas sin mostrar de dónde proviene la información (4, 6). La Tabla 1 resume las opciones disponibles organizadas por etapa del trabajo.

Tabla 1. Herramientas de inteligencia artificial según etapa del trabajo científico biomédico

Etapa del trabajo	Herramientas o recursos	Nivel de respaldo / tipo de fuente	Utilidad práctica	Precaución principal
Formulación de ideas	ChatGPT, Claude, Gemini, Copilot, DeepSeek	Modelos conversacionales; literatura editorial y metodológica en expansión	Explorar enfoques, ordenar ideas, formular preguntas preliminares	No sustituye la pertinencia clínica ni metodológica del investigador; tendencia a confirmar las hipótesis planteadas
Búsqueda biomédica formal	PubMed, Cochrane Library, Embase, Scopus, Web of Science	Bases bibliográficas consolidadas	Identificar literatura trazable y reproducible	Documentar la estrategia cuando se trate de una revisión sistemática
Búsqueda semántica y exploratoria	Semantic Scholar, Elicit, Consensus, Perplexity, OpenRead	Herramientas semánticas; respaldo variable según la herramienta	Buscar por preguntas, resumir resultados iniciales, identificar artículos clave	Verificar siempre en las fuentes originales
Consulta clínica basada en fuentes	OpenEvidence, Vera Health	Documentación institucional o comercial; evaluación independiente variable	Orientar preguntas clínicas y localizar literatura asociada	No citar la respuesta resumida como fuente final ni asumir validación clínica
Mapeo de literatura	ResearchRabbit, Connected Papers, Litmaps, Iris.ai	Herramientas exploratorias de redes bibliográficas	Visualizar redes de artículos, autores y líneas de investigación	Puede favorecer la literatura más citada o en inglés
Contexto de citación	Scite.ai	Herramienta de análisis de citaciones	Revisar si una cita apoya, contrasta o solo menciona un artículo	No reemplaza la lectura crítica del estudio
Cribado de estudios	Rayyan, ASReview, Covidence	Herramientas usadas en revisión de evidencia; algunas con funciones asistidas por IA	Organizar títulos y resúmenes, duplicados y decisiones de inclusión	Requiere revisión humana de los registros dudosos
Extracción preliminar	Elicit, Scholarcy, SciSpace, modelos conversacionales	Evidencia metodológica inicial y heterogénea	Convertir artículos en tablas de diseño, población y resultados	Verificar cada dato contra el PDF original
Gestión bibliográfica	Zotero, Mendeley, EndNote	Gestores bibliográficos consolidados	Insertar citas, ordenar la bibliografía, aplicar Vancouver	No verifica la correspondencia entre la cita y la afirmación
Redacción y estructura	ChatGPT, Claude, Gemini, Copilot	Modelos conversacionales; utilidad editorial con supervisión	Mejorar el flujo, los subtítulos, las transiciones y la claridad	Riesgo de texto genérico o de afirmaciones no sustentadas
Edición lingüística	Grammarly, Writefull, Paperpal, LanguageTool	Herramientas de edición; respaldo variable	Corregir gramática, tono y consistencia	No corrige debilidades científicas
Traducción	DeepL, DeepL Write, modelos conversacionales	Herramientas de traducción y edición lingüística	Preparar versiones iniciales en inglés o español	Revisar la terminología médica y el grado de certeza
Declaración de uso de IA	GAMER, ICMJE, COPE, WAME, normas de la revista	Guías de reporte y documentos editoriales	Transparentar herramienta, uso, verificación y responsabilidad	Adaptar la declaración a cada revista

La inclusión de una herramienta en esta tabla no implica recomendación editorial, validación clínica ni superioridad frente a otras alternativas. Su utilidad depende del objetivo, la trazabilidad de las fuentes, la protección de datos, la evidencia disponible y la verificación humana.

DISCUSIÓN

La IA ya forma parte del ecosistema de investigación biomédica. La proporción documentada por Kobak y cols. (5), con al menos el 13,5 % de los resúmenes de PubMed de 2024 procesados con LLM, sugiere que el debate no debe centrarse solo en una adopción futura, sino también en las condiciones editoriales del uso actual. Su utilidad es clara en las tareas de apoyo, aunque su incorporación no modifica el principio central de la publicación científica: el manuscrito debe ser verificable. Una redacción elegante no equivale a evidencia, del mismo modo que una referencia sugerida por el modelo no garantiza una fuente real ni una tabla bien presentada asegura una extracción correcta de los datos.

La experiencia de quien trabaja con estas herramientas muestra dónde reside el control humano. Obtener un producto de calidad rara vez ocurre al primer intento, pues el resultado útil surge de iterar, de contrastar lo que el modelo entrega con lo que el autor sabe y de reorientar la siguiente solicitud. El pensamiento crítico humano cumple dos funciones distintas en ese ciclo, ya que detecta cuándo la respuesta es insuficiente y decide cómo debe corregirse la instrucción. Esta operación, que consiste en reconocer la falla y guiar la iteración, no puede delegarse en el propio modelo, porque depende del juicio sobre la pregunta, la evidencia esperada y el destinatario del texto. Eso es, en última instancia, lo que separa un manuscrito asistido por IA de un manuscrito sustituido por IA.

Para los autores que publican desde América Latina, esta diferencia tiene consecuencias específicas. El acceso desigual a la literatura indexada, las barreras del inglés académico, los recursos limitados para la asesoría metodológica y el costo de algunas herramientas hacen que la IA pueda funcionar como un apoyo real, sobre todo en la traducción asistida y en la organización de borradores. Sin embargo, esa misma asimetría incrementa el riesgo cuando el autor, presionado por los plazos editoriales, acepta sin verificar lo que el modelo entrega. En este escenario, la formación en alfabetización digital crítica resulta tan importante como el acceso a las propias herramientas. La revista científica latinoamericana tiene aquí un papel doble: exigir una declaración transparente y, al mismo tiempo, no penalizar el uso responsable que reduce barreras estructurales históricas.

El equilibrio razonable está documentado en la literatura editorial reciente (3, 9, 17, 22). La IA no puede figurar como autora, tampoco constituir la fuente primaria de evidencia ni el árbitro final de la calidad metodológica, un punto en el que coinciden WAME (World Association of Medical Editors) (29) y COPE (Committee on Publication Ethics) (30). La declaración GAMER ofrece la estructura más completa disponible hoy para informar el uso de IA (3), aunque no exime al autor de revisar las normas particulares de la revista, dada la heterogeneidad documentada en las políticas editoriales (23). Para el investigador biomédico, la regla práctica se reduce a una división de tareas que la propia naturaleza de los modelos hace evidente: la IA ayuda a encontrar, ordenar, resumir y redactar, mientras que el autor pregunta, lee, verifica, interpreta y responde por el manuscrito.

Esta revisión tiene limitaciones. El campo evoluciona con rapidez y algunas herramientas pueden modificar sus funciones, sus modelos de negocio, sus fuentes de datos o sus políticas de privacidad en plazos breves. Varias aplicaciones disponibles cuentan con una documentación comercial más desarrollada que su evaluación académica independiente, por lo que sus funciones se describen como recursos prácticos y no como intervenciones validadas. No se realizó una revisión sistemática, ya que el propósito fue ofrecer una síntesis narrativa aplicable al trabajo cotidiano del investigador biomédico. Parte de la evidencia disponible procede de publicaciones recientes,

comentarios editoriales, estudios metodológicos iniciales y guías en evolución, de modo que las recomendaciones requieren una actualización periódica.

CONCLUSIONES

La IA generativa puede acompañar al investigador biomédico en casi todas las etapas del trabajo científico, aunque su valor depende de una decisión que sigue siendo humana: identificar cuándo la salida del modelo es suficiente y cuándo requiere reorientación. El uso productivo de estas herramientas es iterativo por naturaleza, y rara vez la primera respuesta resuelve la pregunta. El pensamiento crítico del autor, que reconoce las debilidades de cada borrador y guía la siguiente instrucción, es lo que convierte una sucesión de respuestas plausibles en un manuscrito verificable.

En términos prácticos, esto se traduce en una recomendación concreta para quien publica en el ámbito biomédico: tratar a la IA como un colaborador metodológicamente subordinado y no como un autor delegado. Conviene verificar cada referencia en PubMed o Crossref antes de incluirla, contrastar cada afirmación factual con la fuente original, evitar la introducción de datos identificables o manuscritos confidenciales en plataformas públicas y declarar el uso de IA de forma específica, con indicación de la herramienta, la función y la verificación realizada.

La pregunta abierta para los próximos años no será si la IA participa en la producción científica, algo que los datos ya confirman (5), sino cómo distinguir editorialmente el uso que mejora la calidad del manuscrito del que la deteriora. Esa distinción dependerá de tres condiciones: declaraciones más precisas que las actuales, formación de los autores en la evaluación crítica de las salidas de IA y guías editoriales que se actualicen al ritmo de la tecnología. Mientras esos elementos se consolidan, el principio que sostiene la revisión por pares no cambia, pues la responsabilidad intelectual del manuscrito sigue siendo, de forma indelegable, del autor humano que lo firma.

AGRADECIMIENTOS

La autora agradece a la Dra. Mireya González Blanco por su apoyo y orientación editorial durante la preparación del manuscrito.

FINANCIAMIENTO

la autora declara no haber recibido financiamiento específico de agencias públicas, comerciales o sin fines de lucro.

La autora declara no tener conflictos de interés.

DECLARACIÓN DE USO DE INTELIGENCIA ARTIFICIAL

Durante la preparación de este manuscrito se emplearon, como apoyo, los modelos más recientes disponibles de OpenAI (ChatGPT 5.5), Google (Gemini 3.5) y Anthropic (Claude Fable 5). Su uso se limitó a la organización temática del artículo y a la revisión gramatical y de estilo del texto en español e inglés. Ninguna de estas herramientas generó datos, resultados ni conclusiones. Todas las sugerencias fueron revisadas por la autora, contrastadas con las fuentes originales y modificadas cuando resultó necesario. No se introdujo en dichas

herramientas información clínica identificable ni datos confidenciales de pacientes. La autora asume la responsabilidad plena por el contenido, las citas y las conclusiones del manuscrito.

REFERENCIAS

1. Yoga Ratnam KK. Generative artificial intelligence in public health research and scientific communication: a narrative review of real applications and future directions. Digit Health. 2025;11:1-10. DOI:10.1177/20552076251362070.
2. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29(8):1930-1940. DOI:10.1038/s41591-023-02448-8.
3. 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 Evid Based Med. 2025;30(6):390-400. DOI:10.1136/bmjebm-2025-113825.
4. Jin Q, Leaman R, Lu Z. PubMed and beyond: biomedical literature search in the age of artificial intelligence. EBioMedicine. 2024;100:104988. DOI: 10.1016/j.ebiom.2024.104988.
5. Kobak D, González-Márquez R, Horvát EÁ, Lause J. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. Sci Adv. 2025;11(27):eadt3813. DOI: 10.1126/sciadv.adt3813.
6. Chelli M, Descamps J, Lavoué V, Trojani C, Azar M, Deckert M, *et al*. Hallucination rates and reference accuracy of ChatGPT and Bard for systematic reviews: comparative analysis. J Med Internet Res. 2024;26:e53164. DOI:10.2196/53164.
7. Salvagno M, Taccone FS, Gerli AG. Can artificial intelligence help for scientific writing? Crit Care. 2023;27:75. DOI:10.1186/s13054-023-04380-2.
8. Sen CK. Artificial intelligence tools in biomedical research: part 1—literature search and knowledge mining. Antioxid Redox Signal. 2026;44(1-3):1-10. DOI: 10.1177/15230864251405885.
9. Bernard N, Sagawa Y Jr, Bier N, Lihoreau T, Pazart L, Tannou T. Using artificial intelligence for systematic review: the example of Elicit. BMC Med Res Methodol. 2025;25(1):75. DOI:10.1186/s12874-025-02528-y.
10. OpenEvidence. The leading medical information platform [Internet]. Cambridge: OpenEvidence; s/f [consultado 30 de abril de 2026]. Disponible en: https://www.openevidence.com/
11. Vera Health. Evidence-based clinical answers [Internet]. San Francisco: Vera Health; s/f [consultado 30 de abril de 2026]. Disponible en: https://www.verahealth.ai/
12. Luo X, Chen F, Zhu D, Wang L, Wang Z, Liu H, *et al*. Potential roles of large language models in the production of systematic reviews and meta-analyses. J Med Internet Res. 2024;26:e56780. DOI:10.2196/56780.
13. Gwon YN, Kim JH, Chung HS, Jung EJ, Chun J, Lee S, *et al*. The use of generative AI for scientific literature searches for systematic reviews: ChatGPT and Microsoft Bing AI performance evaluation. JMIR Med Inform. 2024;12:e51187. DOI:10.2196/51187.
14. Gartlehner G, Kahwati L, Hilscher R, Thomas I, Kugley S, Crotty K, *et al*. Data extraction for evidence synthesis using a large language model: a proof-of-concept study. Res Synth Methods. 2024;15(4):576-589. DOI:10.1002/jrsm.1710.
15. 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. DOI:10.2196/48996.
16. Hasan B, Saadi S, Rajjoub NS, Hegazi M, Al-Kordi M, Fleti F, *et al*. Integrating large language models in systematic reviews: a framework and case study using ROBINS-I for risk of bias assessment. BMJ Evid Based Med. 2024;29(6):394-398. DOI:10.1136/bmjebm-2023-112597.
17. Leung TI, de Azevedo Cardoso T, Mavragani A, Eysenbach G. Best practices for using AI tools as an author, peer reviewer, or editor. J Med Internet Res. 2023;25:e51584. DOI: 10.2196/51584.
18. Hwang SI, Lim JS, Lee RW, Matsui Y, Iguchi T, Hiraki T, *et al*. Is ChatGPT a "Fire of Prometheus" for non-native English-speaking researchers in academic writing? Korean J Radiol. 2023;24(10):952-959. DOI:10.3348/kjr.2023.0773.
19. Del Giglio A, da Costa MUP. The use of artificial intelligence to improve the scientific writing of non-native English speakers. Rev Assoc Med Bras (1992). 2023;69(9):e20230560. DOI:10.1590/1806-9282.20230560.
20. Carneiro MM. Artificial intelligence in scientific writing: sailing fair winds or between the devil and the deep blue sea? Women Health. 2025;65(1):1-3. DOI:10.1080/03630242.2025.2445890.
21. Chetwynd E. Ethical use of artificial intelligence for scientific writing: current trends. J Hum Lact. 2024;40(2):211-215. DOI:10.1177/08903344241235160.
22. Flanagin A, Kendall-Taylor J, Bibbins-Domingo K. Guidance for authors, peer reviewers, and editors on use of AI, language models, and chatbots. JAMA. 2023;330(8):702-703. DOI:10.1001/jama.2023.12500.

23. Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, *et al*. Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. BMJ. 2024;384:e077192. DOI:10.1136/bmj-2023-077192.
24. Yoo JH. Defining the boundaries of AI use in scientific writing: a comparative review of editorial policies. J Korean Med Sci. 2025;40(23):e187. DOI:10.3346/jkms.2025.40.e187.
25. Yin S, Huang S, Xue P, Xu Z, Lian Z, Ye C, *et al*. Generative artificial intelligence usage guidelines for scholarly publishing: a cross-sectional study of medical journals. BMC Med. 2025;23:77. DOI:10.1186/s12916-025-03899-1.
26. International Committee of Medical Journal Editors. Artificial intelligence: AI use by authors [Internet]. ICMJE; N/D [consultado 30 de abril de 2026]. Disponible en: https://www.icmje.org/recommendations/browse/artificial-intelligence/ai-use-by-authors.html
27. Li H, Moon JT, Purkayastha S, Celi LA, Trivedi H, Gichoya JW. Ethics of large language models in medicine and medical research. Lancet Digit Health. 2023;5(6):e333-e335. DOI:10.1016/S2589-7500(23)00083-3.
28. Ramoni D, Sgura C, Liberale L, Montecucco F, Ioannidis JPA, Carbone F. Artificial intelligence in scientific medical writing: legitimate and deceptive uses and ethical concerns. Eur J Intern Med. 2024;127:31-35. DOI:10.1016/j.ejim.2024.07.012.
29. Zielinski C, Winker M, Aggarwal R, Ferris L, Heinemann M, Lapeña JF Jr, *et al*. Chatbots, ChatGPT, and scholarly manuscripts: WAME recommendations on ChatGPT and chatbots in relation to scholarly publications. Natl Med J India. 2023;36(1):1-4. DOI:10.25259/NMJI_365_23.
30. Committee on Publication Ethics. Authorship and AI tools [Internet]. COPE; 2023 [consultado 30 de abril de 2026]. Disponible en: https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools

Recibido 11 de mayo de 2026
Aprobado para publicación 5 de junio de 2026