

Chapter 6 / Capítulo 6

Critical reading of scientific articles on
health: A practical guide (Spanish Edition)
ISBN: 978-9915-9928-3-9
DOI: 10.62486/978-9915-9928-3-9.Ch06
Pages: 51-69



© 2026 The authors. This is an open access
book chapter distributed under the terms of
the Creative Commons Attribution (CC BY)
4.0 License.

Búsqueda de artículos de investigación con inteligencia artificial

Searching for Research Articles with Artificial Intelligence

Mauricio Alejandro Paz Del Rio  

Blas Apaza-Huanca  

1 Clínica Los Andes. La Paz, Bolivia

2 Ministerio de Salud y Deportes. Instituto Académico Científico Quispe-Cornejo. La Paz, Bolivia.

ABSTRACT

Introduction: Artificial intelligence (AI) has transformed the search for scientific literature by optimizing the retrieval, organization, and synthesis of biomedical evidence. Integrated with the principles of Evidence-Based Medicine, it facilitates the formulation of clinical questions using the PICO format, the identification of MeSH/DeCS terms, and the design of reproducible search strategies. Development: This chapter analyzes the use of language models and specialized tools such as Perplexity, Elicit, Scite AI, Consensus, Semantic Scholar, ResearchRabbit, and OpenEvidence to improve the location and preliminary appraisal of scientific studies. It also emphasizes the construction of structured prompts, mandatory reference validation, and the integration of AI with traditional biomedical databases to ensure methodological rigor and traceability of evidence. Conclusions: AI constitutes a methodological assistant that increases the efficiency of the literature search process, but it does not replace clinical judgment, critical appraisal, or the verification of original sources. From a reflective perspective, its responsible use requires methodological skills, critical thinking, and ethical commitment to ensure clinical decisions are based on reliable, up-to-date, and reproducible scientific evidence.

Keywords: Literature search; Artificial intelligence; PICO format; MeSH descriptors; Apocryphal citation; Active systematic review

RESUMEN

Introducción: La inteligencia artificial (IA) ha transformado la búsqueda de literatura científica al optimizar la recuperación, organización y síntesis de la evidencia biomédica. Integrada con los principios de la Medicina Basada en la Evidencia, facilita la formulación de preguntas clínicas mediante el formato PICO, la identificación de términos MeSH/DeCS y el diseño de estrategias de búsqueda reproducibles. Desarrollo: El capítulo analiza el empleo de modelos de lenguaje y herramientas especializadas como Perplexity, Elicit, Scite AI, Consensus, Semantic Scholar, Research Rabbit y OpenEvidence para mejorar la localización y evaluación preliminar de estudios científicos. Asimismo, enfatiza la construcción de prompts estructurados, la validación obligatoria de referencias y la integración de la IA con bases de datos biomédicas tradicionales para garantizar rigor metodológico y trazabilidad de la evidencia. Conclusiones: La IA constituye un asistente metodológico que incrementa la eficiencia del proceso de búsqueda bibliográfica, pero no sustituye el juicio clínico, la lectura crítica ni la verificación de las fuentes originales. Desde una perspectiva reflexiva, su utilización responsable exige competencias metodológicas, pensamiento crítico y compromiso ético para asegurar decisiones clínicas fundamentadas en evidencia científica confiable, actualizada y reproducible.

Palabras clave: Búsqueda bibliográfica; Inteligencia artificial; Formato PICO; Descriptores MeSH; Cita apócrifa; Revisión sistemática viva

1. CASO CLÍNICO

Un residente de Medicina Interna atiende a un paciente masculino de 58 años con hipertensión arterial de difícil control, obesidad y apnea obstructiva del sueño (AOS). El médico sospecha de hiperalдостерonismo primario y necesita la mejor evidencia disponible sobre: prevalencia en pacientes con AOS. pruebas diagnósticas más exactas, algoritmos diagnósticos recientes. En el hospital, el tiempo es limitado y las búsquedas tradicionales en PubMed necesitan tiempo para resultados exactos. El residente decide utilizar herramientas de inteligencia artificial para acelerar la búsqueda y filtrar información científica de calidad.

2. INTRODUCCIÓN

La búsqueda eficiente de información científica es un pilar de la Medicina Basada en la Evidencia (MBE). Tradicionalmente, se ha dependido de bases de datos como PubMed, Embase, SciELO o la Biblioteca Virtual en Salud (BVS), utilizando estrategias de búsqueda con términos MeSH, DeCS y operadores booleanos. Sin embargo, en la era digital, la inteligencia artificial (IA) ofrece nuevas herramientas que agilizan la búsqueda, resumen la evidencia y permiten formular preguntas más precisas, combinando el rigor de las bases de datos con la rapidez del procesamiento de lenguaje natural.

Estas herramientas no reemplazan la búsqueda manual ni la evaluación crítica del artículo, pero pueden ahorrar tiempo, sugerir términos de búsqueda más eficientes, organizar la evidencia y poder comunicarse con el mismo artículo si se tiene dudas.

3. TÉRMINOS MESH, DECS Y OPERADORES BOOLEANOS CON LLM

En la búsqueda de literatura médica, el uso de términos controlados es fundamental para recuperar artículos relevantes y evitar pérdidas de información. En medicina, dos sistemas destacan por su uso estandarizado:

MeSH (Medical Subject Headings): es el tesoro desarrollado por la Biblioteca Nacional de Medicina de los Estados Unidos, utilizado por PubMed/MEDLINE para clasificar artículos bajo descriptores estandarizados.

DeCS (Descriptores en Ciencias de la Salud): es la adaptación de MeSH al español y portugués, desarrollada por BIREME/OPS/OMS, utilizada en LILACS y en la BVS.

La importancia de estos términos radica en que los autores no siempre utilizan la misma palabra para describir un concepto. Por ejemplo, un estudio sobre “ataque cardíaco” puede estar indexado bajo el descriptor “Infarto de Miocardio”. Si solo se busca “ataque cardíaco” en el texto libre, se podrían perder artículos relevantes. Para optimizar la recuperación de evidencia, los términos MeSH y DeCS suelen combinarse con operadores booleanos, estas herramientas constituyen el lenguaje formal de la búsqueda científica y permiten establecer relaciones lógicas entre conceptos:

- AND, para combinar conceptos distintos y restringir la búsqueda.
- OR, para agrupar sinónimos o términos equivalentes y ampliar la búsqueda.
- NOT, para excluir términos no deseados, con uso cauteloso.
- Paréntesis y comillas: permite controlar el orden lógico de la búsqueda y recuperar frases exactas, respectivamente.

Tradicionalmente, la identificación de descriptores adecuados y su correcta combinación requería experiencia y tiempo (como se vio en el capítulo 5). En la actualidad, los modelos de lenguaje grande (LLM), como ChatGPT, Gemini y otras inteligencias artificiales conversacionales, pueden facilitar este proceso al:

- Sugerir rápidamente términos MeSH y DeCS pertinentes.
- Proponer sinónimos y términos de texto libre utilizados en la literatura.
- Construir estrategias preliminares de búsqueda adaptadas a distintas bases de datos.

No obstante, es importante recordar que los LLM no sustituyen la comprensión metodológica del investigador, sino que actúan como herramientas de apoyo para acelerar y optimizar la búsqueda.

4. FORMATO PICO CON LLM

El formato PICO (Paciente, Intervención, Comparador, Outcome) es el más utilizado porque organiza los elementos clave de la pregunta de forma estructurada. PICO significa:

- **P (Paciente o Problema):** grupo, condición o población que nos interesa.
- **I (Intervención o Exposición):** tratamiento, prueba diagnóstica o factor de riesgo que queremos investigar.
- **C (Comparador):** alternativa frente a la cual comparamos la intervención. Puede ser placebo, estándar de cuidado o ausencia de exposición.
- **O (Outcome o Desenlace):** resultado clínico o biomédico que queremos medir.

La solicitud precisa de esto en los LLM nos brinda muchas ventajas: Ahorra tiempo en la estructuración inicial de la pregunta, puede generar variantes de la pregunta adaptadas a diagnósticos, terapias o pronóstico, sugiere palabras clave para usar en la búsqueda, puede formular la pregunta en distintos idiomas.

5. PREGUNTA DE INVESTIGACIÓN CON LLM

Formular una pregunta de investigación bien estructurada constituye un paso central en la MBE, ya que determina el enfoque de la búsqueda bibliográfica, la selección de los estudios relevantes y la posterior interpretación crítica de la literatura. Una pregunta mal definida conduce, con frecuencia, a búsquedas imprecisas, resultados heterogéneos y conclusiones de limitada aplicabilidad clínica.

En este contexto, el formato PICO es la herramienta ampliamente utilizada para organizar la pregunta clínica de manera sistemática. Su valor no radica únicamente en facilitar una búsqueda más eficiente, sino en obligar al clínico o investigador a explicitar desde el inicio los elementos clave del problema, promoviendo un análisis estructurado y orientado a la toma de decisiones.

Los LLM pueden apoyar este proceso al actuar como asistentes metodológicos. Su utilidad principal reside en ayudar a descomponer un problema clínico complejo en sus componentes PICO, proponer formulaciones alternativas de la pregunta de investigación y sugerir términos clave asociados a cada componente. De este modo, los LLM contribuyen a reducir errores frecuentes, como preguntas excesivamente amplias o ambiguas, sin reemplazar el razonamiento clínico ni la evaluación crítica del investigador.

Es importante destacar que el uso de LLM en esta etapa debe entenderse como un apoyo estructural y no como una automatización del proceso intelectual. La validez y relevancia de la pregunta de investigación siguen dependiendo del juicio clínico, del contexto asistencial y del conocimiento previo del investigador.

6. CONSTRUCCIÓN DE UN PROMPT PARA LA BÚSQUEDA DE LITERATURA CIENTÍFICA CON INTELIGENCIA ARTIFICIAL

Uno de los errores más frecuentes al utilizar inteligencia artificial en la búsqueda de artículos científicos es formular instrucciones demasiado breves o inespecíficas. Expresiones como “dame los términos MeSH de hiperalдостерonismo” suelen producir respuestas incompletas, desordenadas o poco útiles para una búsqueda real en bases de datos biomédicas.

En investigación en salud, la calidad de los resultados obtenidos mediante inteligencia artificial depende directamente de la calidad del prompt, es decir, de la instrucción que se le proporciona a la herramienta. En este sentido, el prompt cumple una función metodológica similar a la de una pregunta PICO bien formulada: orienta el proceso, delimita el alcance y mejora la reproducibilidad de la búsqueda. Para que la IA aporte verdadero valor a la MBE, el prompt debe construirse de forma estructurada, explícita y orientada a objetivos científicos, no como una consulta informal.

6.1 PRINCIPIOS PARA UN PROMPT CIENTÍFICAMENTE SÓLIDO.

Un prompt adecuado para búsqueda bibliográfica debe integrar varios elementos clave.

En primer lugar, es fundamental definir el rol que debe asumir la IA. Cuando no se especifica, la IA tiende a responder desde un enfoque general o divulgativo. En el contexto de la investigación clínica, se debe indicar que actúe como experto en búsqueda de literatura biomédica y en el uso

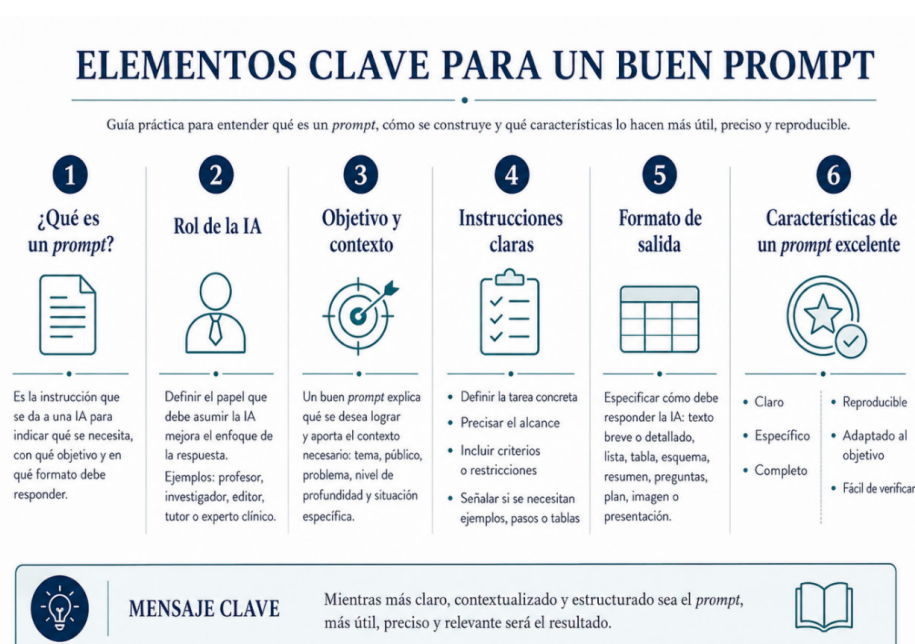
de tesauros controlados, como MeSH y DeCS.

En segundo lugar, debe establecerse con claridad el **objetivo de la tarea**. El propósito no es aprender qué son los términos MeSH, sino optimizar una búsqueda real en bases de datos científicas como PubMed y la BVS, con fines clínicos o investigativos.

El tercer elemento es el **contexto clínico**, que debe ser específico y bien delimitado. Incluir la condición de interés, la población (por ejemplo, adultos o pediátricos) y el enfoque clínico evita ambigüedades y mejora la selección de descriptores. Un contexto mal definido puede conducir a resultados irrelevantes o excesivamente amplios.

En cuarto lugar, el prompt debe describir de manera explícita la **tarea concreta** que se espera que la IA realice. En una búsqueda bibliográfica rigurosa, esto implica solicitar la identificación de términos MeSH oficiales, sus equivalentes en DeCS, sinónimos y términos de texto libre, así como la construcción de estrategias de búsqueda con operadores booleanos. Además, es recomendable especificar el tipo de estudios de interés (por ejemplo, prevalencia, diagnóstico, estudios observacionales) y el periodo temporal prioritario, especialmente en un contexto de evidencia actualizada.

Figura 1. Elementos clave para un buen prompt



Fuente: Elaboración propia con apoyo de Inteligencia artificial

Es esencial indicar el **formato de salida**. En investigación, la información estructurada en tablas facilita la lectura, la comparación y la reproducibilidad. Solicitar desde el inicio una salida tabular transforma la respuesta de la IA en una herramienta práctica, lista para ser utilizada o documentada en un protocolo de investigación.

6.2 PROMPT INTEGRAL PARA ESTRATEGIA DE BÚSQUEDA Y FORMULACIÓN DE LA PREGUNTA DE INVESTIGACIÓN CON INTELIGENCIA ARTIFICIAL

Una vez comprendidos los principios para construir prompts de calidad, es posible integrar en una sola instrucción todos los componentes clave del proceso de búsqueda en MBE: la formulación de la pregunta de investigación, la estructuración PICO, la identificación de términos MeSH y DeCS, y la construcción de estrategias de búsqueda reproducibles en bases de datos biomédicas.

Este enfoque integral permite que la IA funcione como un asistente metodológico, capaz de acompañar al investigador desde el planteamiento del problema clínico hasta la ejecución de la búsqueda bibliográfica, sin sustituir el razonamiento crítico ni la validación humana. A continuación, se presenta un modelo de prompt completo, diseñado para ser utilizado de forma

reproducible en distintos contextos clínicos.

Ejemplo de Prompt:

Actúa como un experto en búsqueda de literatura biomédica, Medicina Basada en la Evidencia y metodología de investigación clínica.

Objetivo: ayudarme a formular una pregunta de investigación clara y a diseñar una estrategia de búsqueda rigurosa, reproducible y optimizada para PubMed y la Biblioteca Virtual en Salud, con fines de investigación clínica.

Contexto clínico:

Hiperaldosteronismo primario en pacientes adultos con apnea obstructiva del sueño enfocado a estudios de prevalencia y estudios de exactitud diagnóstica.

Tareas específicas:

- Formular la pregunta de investigación utilizando el enfoque PICO.
- Desglosar explícitamente los componentes P (Paciente), I (Intervención o exposición), C (Comparador) y O (Outcome).
- Proponer al menos dos variantes completas de la pregunta PICO, orientadas a prevalencia y diagnóstico.
- Identificar únicamente términos MeSH oficiales utilizados en PubMed y sus equivalentes DeCS validados en español.
- Incluir sinónimos, acrónimos y términos de texto libre comúnmente utilizados por los autores en los artículos científicos.
- Proponer una estrategia de búsqueda estructurada con operadores booleanos optimizada para PubMed y otra para la BVS, listas para ser copiadas y ejecutadas.
- Indicar, cuando sea posible, el tipo de estudios que cada término suele recuperar (por ejemplo: estudios de prevalencia, estudios observacionales, estudios de exactitud diagnóstica).

Formato de salida:

Presenta la información de manera estructurada en tablas, incluyendo: Tabla PICO (P, I, C, O, términos MeSH/DeCS asociados). Tabla de descriptores con las siguientes columnas:

- Descriptor MeSH
- Descriptor DeCS
- Sinónimos o términos de texto libre
- Tipo de estudio asociado
- Estrategia de búsqueda sugerida para PubMed y BVS.

Figura2. Prompt para estrategia de búsqueda y formulación de la pregunta de investigación con Inteligencia artificial



Fuente: Elaboración propia con apoyo de inteligencia artificial

Este tipo de prompt no solo mejora la calidad y pertinencia de los resultados obtenidos, sino que también forma al lector en pensamiento metodológico, permitiéndole comprender que la inteligencia artificial no reemplaza al investigador. Por el contrario, actúa como una herramienta que amplifica su capacidad para formular preguntas relevantes, estructurar búsquedas complejas y filtrar evidencia científica de manera eficiente. En el contexto de la MBE, aprender a construir prompts de este nivel constituye una competencia metodológica esencial, comparable a saber formular una pregunta PICO o diseñar una estrategia de búsqueda avanzada en PubMed.

6.3 PROMPT INTEGRAL PARA BÚSQUEDA DE ARTÍCULOS DE INVESTIGACIÓN CON LLM

Además de ayudar a formular PICO, identificar MeSH/DeCS y construir cadenas booleanas, los LLM pueden emplearse de manera más “operativa”: pedirles que propongan directamente un conjunto inicial de artículos relevantes a partir de una pregunta clínica o de investigación. En este rol, el LLM actúa como un asistente de cribado temprano, que sugiere literatura probable (revisiones sistemáticas, metaanálisis, ensayos clínicos, estudios de cohorte, etc.) y organiza la información en un formato útil para lectura rápida. La ventaja principal es pragmática: en escenarios clínicos con limitación de tiempo, un LLM puede acelerar tres tareas que suelen ser lentas cuando se realizan manualmente:

- Priorización inicial de evidencia: sugiere tipos de estudio y fuentes que suelen ofrecer mayor valor (p. ej., revisiones sistemáticas recientes y ensayos clínicos clave).
- Extracción de información “de primera pasada”: resume pregunta, población, comparadores y desenlaces principales, permitiendo decidir rápidamente qué artículos vale la pena leer completos.
- Organización comparativa: presenta los estudios en tabla y facilita comparar resultados (eficacia, seguridad, subgrupos, limitaciones).

Sin embargo, este uso directo exige una advertencia metodológica: un LLM puede equivocarse en detalles bibliográficos (autores, año, DOI) o proponer artículos que no responden exactamente a la pregunta. Por tanto, el uso correcto en MBE es: LLM para generar una lista inicial más verificación obligatoria en la base de datos (PubMed/Scopus/Cochrane/otros) antes de citar o concluir. Dicho de otro modo, el LLM acelera el triage, pero la validez final depende de la comprobación en la fuente.

REDACCIÓN DEL PROMPT

Para que esta modalidad funcione, el prompt debe dejar claro que el objetivo no es solo una explicación, sino una curaduría de evidencia con trazabilidad. En particular, sumado a las características descritas previamente sobre el prompt, debe exigir:

- Delimitación explícita: periodo temporal (p. ej., últimos 5 años), población, comparadores, desenlaces.
- Jerarquización de evidencia: priorizar revisiones sistemáticas/metaanálisis, ensayos y luego estudios observacionales de alta calidad.
- Metadatos obligatorios: título, revista, año, tipo de estudio, hallazgos, y enlace/identificador verificable.
- Estructura de salida: tabla más síntesis crítica comparativa.
- Control de calidad: indicar que solo incluya artículos con enlace verificable (PubMed ID/URL, DOI), y que marque cualquier dato incierto.

FORMATO DE PROMPT

Actúa como un médico investigador experto en Medicina Interna, Endocrinología (dependiendo de la rama del trabajo que se está buscando) y Medicina Basada en la Evidencia, con experiencia en búsqueda avanzada y síntesis crítica de literatura científica.

Objetivo: Ayudarme a identificar y sintetizar la mejor evidencia científica disponible para orientar la toma de decisiones clínicas y diagnósticas.

Contexto clínico: Paciente masculino de 58 años con hipertensión arterial de difícil control, obesidad y apnea obstructiva del sueño. Se sospecha hiperaldosteronismo primario como causa secundaria de hipertensión.

Tareas: Necesito información sobre: Prevalencia del hiperaldosteronismo primario en

pacientes con apnea obstructiva del sueño, pruebas diagnósticas con mayor exactitud para el diagnóstico de hiperaldosteronismo primario en este contexto, algoritmos diagnósticos recientes que integren hipertensión, apnea obstructiva del sueño e hiperaldosteronismo primario.

INSTRUCCIONES ESPECÍFICAS:

- Realiza una búsqueda dirigida de la mejor evidencia publicada en los últimos 5 a 10 años, priorizando revisiones sistemáticas, metaanálisis, ensayos clínicos y estudios observacionales de alta calidad.
- Selecciona entre 8 y 10 artículos relevantes provenientes de revistas indexadas en PubMed, Scopus o Cochrane.
- Para cada artículo incluye obligatoriamente:
 - Título del artículo
 - Autor principal
 - Revista y año de publicación
 - Tipo de estudio
 - Principales hallazgos relacionados con prevalencia, exactitud diagnóstica o algoritmos diagnósticos
- Enlace verificable (PMID o DOI).
- Organiza la información en una tabla comparativa para facilitar la lectura.
- Finaliza con una síntesis crítica, integrando los hallazgos y destacando:
 - Magnitud de la prevalencia reportada en pacientes con AOS.
 - Pruebas diagnósticas con mejor rendimiento (sensibilidad, especificidad, valor clínico).
 - Consensos actuales y limitaciones de la evidencia disponible.
 - Indica explícitamente cualquier dato que requiera verificación adicional.

FORMATO DE SALIDA:

Figura 3. Prompt para búsqueda de artículos científicos con Inteligencia Artificial



Fuente: Elaboración propia con apoyo de inteligencia artificial

Tabla comparativa con las siguientes columnas: Título del artículo. Autor principal. Revista (año). Tipo de estudio. Población estudiada. Hallazgos clave (prevalencia / pruebas diagnósticas / algoritmo). Enlace verificable (PMID o DOI). Síntesis crítica final, estructurada en párrafos breves, que incluya:

- Magnitud de la prevalencia del hiperaldosteronismo primario reportada en pacientes con

apnea obstructiva del sueño.

- Pruebas diagnósticas con mejor rendimiento (sensibilidad, especificidad y valor clínico).
- Consensos actuales sobre algoritmos diagnósticos y puntos de controversia.
- Limitaciones metodológicas de la evidencia disponible y brechas de conocimiento.

Nota de verificación: Indica explícitamente cualquier dato bibliográfico, resultado o afirmación que requiera verificación adicional en la fuente original.

6. USO DE OTRAS IA PARA ARTÍCULOS DE INVESTIGACIÓN

En los últimos años, han surgido diversas herramientas de IA diseñadas específicamente para agilizar la búsqueda, organización y síntesis de información científica. A diferencia de buscadores generales, estas plataformas están optimizadas para trabajar con literatura revisada por pares, bases de datos académicas y análisis de citaciones. En este acápite, analizaremos las más relevantes aplicadas a nuestro caso clínico y pregunta PICO.

A. PERPLEXITY AI

Perplexity integra un motor de búsqueda con un modelo de lenguaje que responde en formato conversacional y, de manera característica, cita explícitamente las fuentes consultadas. Esta funcionalidad lo convierte en una herramienta especialmente útil para obtener una visión rápida y orientativa del estado del arte, identificar revisiones recientes y acceder de forma directa a enlaces de PubMed y otros repositorios académicos.

En investigación clínica, Perplexity funciona mejor como herramienta de exploración inicial. Permite reconocer con rapidez qué líneas de evidencia existen, cuáles son los autores o grupos más citados y qué artículos actúan como referencias centrales del tema. No obstante, sus resultados deben considerarse un punto de partida, ya que la búsqueda final, exhaustiva y reproducible debe realizarse directamente en las bases de datos biomédicas. Es especialmente útil para orientarse rápidamente, identificar revisiones recientes y localizar artículos “semilla” que luego pueden ampliarse mediante búsqueda clásica y revisión de referencias. No debe utilizarse como único método cuando el objetivo es una revisión exhaustiva o sistemática.

Aplicación al caso clínico: Si se introduce una pregunta derivada del problema planteado al inicio del capítulo, por ejemplo:

“En adultos con apnea obstructiva del sueño, ¿cuál es la prevalencia de hiperaldosteronismo primario en comparación con adultos sin apnea obstructiva del sueño?”

Perplexity suele devolver un resumen narrativo acompañado de enlaces a PubMed y otras fuentes académicas. Una estrategia recomendable es solicitar explícitamente academic sources y pedir que los resultados se ordenen por fecha o relevancia, según el objetivo de la búsqueda.

Tabla 1 Ventajas y limitaciones de Perplexity AI para la búsqueda inicial de literatura científica

Ventajas	Limitaciones
Resultados rápidos, útiles en contextos clínicos con tiempo limitado.	No garantiza una búsqueda exhaustiva ni sistemática.
Enlaces directos a fuentes académicas (PubMed, editoriales).	Puede omitir artículos clave o relevantes.
Permite refinar la búsqueda mediante preguntas sucesivas.	Tiende a priorizar estudios más visibles o citados.
Facilita una visión global del estado del arte.	No evalúa la calidad metodológica de los estudios.
Útil para identificar artículos “semilla”.	Requiere verificación manual de datos y resultados.

Fuente: Elaboración propia con apoyo de inteligencia artificial

B. ELICIT

Elicit es una herramienta de inteligencia artificial diseñada específicamente para apoyar la búsqueda y síntesis de literatura científica, con un enfoque claro en investigación en salud y ciencias biomédicas. A diferencia de motores de búsqueda generales, Elicit trabaja directamente sobre bases de datos académicas (principalmente PubMed y Semantic Scholar) y estructura la

información en tablas, lo que facilita el análisis comparativo.

En investigación clínica, Elicit resulta especialmente útil cuando el objetivo es identificar estudios relevantes de manera estructurada, extraer información clave de cada artículo y organizarla de forma sistemática. Su fortaleza principal radica en la capacidad de descomponer preguntas de investigación y vincularlas con evidencia primaria, reduciendo el tiempo invertido en la lectura inicial de múltiples resúmenes.

Al aplicar la pregunta mencionada del caso clínico, Elicit puede recuperar estudios observacionales, revisiones y ensayos relevantes, mostrando para cada uno información estructurada como población, variables evaluadas y principales resultados. Esto permite al personal de salud identificar rápidamente qué artículos ameritan una lectura completa y cuáles no aportan información directa a la pregunta clínica.

Se recomienda usarla cuando se necesita organizar evidencia primaria de forma estructurada, por ejemplo, en la fase inicial de un protocolo de investigación o al comparar resultados entre estudios observacionales. No sustituye la búsqueda exhaustiva en PubMed ni la lectura crítica completa, pero constituye un excelente complemento para priorizar artículos relevantes.

Tabla 2 Ventajas y limitaciones de Elicit para la identificación y organización inicial de estudios científicos

Ventajas	Limitaciones
Orientada específicamente a literatura científica.	No cubre todas las revistas ni bases de datos existentes.
Presenta resultados en tablas estructuradas.	Puede omitir artículos recientes aún no indexados.
Facilita la comparación rápida entre estudios.	Depende de la calidad de los metadatos disponibles.
Útil para identificar estudios observacionales y revisiones.	Menor utilidad para guías clínicas o documentos de consenso.
Ahorra tiempo en la lectura inicial de resúmenes.	Requiere verificación manual de resultados y conclusiones.

Fuente: Elaboración propia con apoyo de inteligencia artificial

C. SCITE AI

Scite AI es una herramienta de inteligencia artificial orientada a un aspecto específico y crítico de la investigación científica: cómo se citan los artículos en la literatura. A diferencia de otras plataformas que priorizan la recuperación o el resumen de estudios, Scite analiza el contexto de las citaciones y las clasifica según su función dentro del artículo citante. Las categoriza las citas como:

- Supporting (citan el artículo para respaldar un hallazgo),
- Contrasting (lo citan para cuestionar o contradecir resultados),
- Mentioning (lo mencionan sin posicionarse críticamente).

Esta característica aporta un valor metodológico importante, ya que permite al investigador identificar no solo qué artículos son citados con frecuencia, sino cómo es utilizada esa evidencia dentro del cuerpo de la literatura científica.

Al aplicarlo en el caso clínico, esta IA puede utilizarse después de identificar artículos clave (por ejemplo, estudios de prevalencia o exactitud diagnóstica). El personal de salud puede analizar si dichos artículos son mayoritariamente citados como evidencia de apoyo o si existen estudios que cuestionan sus conclusiones, lo cual ayuda a valorar la consistencia y robustez de los hallazgos.

Se recomienda utilizarlo después de haber identificado artículos relevantes mediante otras herramientas o búsquedas clásicas. Su mayor valor reside en apoyar la lectura crítica, ayudando a discernir si la evidencia es consistente, controvertida o débilmente sustentada. No debe utilizarse como única herramienta de búsqueda, sino como complemento para evaluar la solidez del cuerpo de evidencia.

Tabla 3 Ventajas y limitaciones de Scite AI para el análisis del contexto de citación en literatura científica

Ventajas	Limitaciones
Analiza el contexto de las citaciones, no solo su número.	No es un motor de búsqueda primario de literatura.
Permite identificar rápidamente evidencia apoyada o controvertida.	Depende de la disponibilidad de textos completos indexados.
Útil para detectar artículos influyentes y debates científicos.	No evalúa directamente la calidad metodológica del estudio.
Facilita la lectura crítica al mostrar citas contrastantes.	Menor utilidad en temas con poca literatura publicada.
Aporta valor en fases avanzadas de análisis.	Requiere complementar con lectura directa del artículo.

Fuente: Elaboración propia con apoyo de inteligencia artificial

D. CONSENSUS

Consensus es una herramienta de inteligencia artificial orientada a responder preguntas clínicas o de investigación basándose exclusivamente en literatura científica revisada por pares. A diferencia de motores de búsqueda generales, Consensus no se apoya en contenido web amplio, sino que extrae respuestas a partir de artículos académicos, con énfasis en revisiones sistemáticas, metaanálisis y ensayos clínicos.

Su principal característica es que intenta sintetizar una respuesta basada en el “consenso” de la evidencia disponible, mostrando si los estudios tienden a apoyar, rechazar o no ser concluyentes respecto a una afirmación específica. Esto la convierte en una herramienta particularmente atractiva para clínicos que necesitan una respuesta rápida y basada en evidencia, sin perder el anclaje científico.

Ante la aplicación al caso clínico este puede ofrecer una respuesta resumida indicando si la mayoría de los estudios disponibles apoyan dicha asociación, junto con enlaces a los artículos que sustentan esa conclusión. Para el personal de salud, esto permite identificar rápidamente si existe respaldo científico consistente para la sospecha clínica y cuáles son los estudios clave que fundamentan dicha relación.

Se recomienda Consensus cuando se necesita una respuesta rápida y orientativa, basada en evidencia científica, para confirmar si una hipótesis clínica cuenta con respaldo en la literatura. Resulta ideal en etapas tempranas de toma de decisiones o como complemento a la lectura crítica, pero no debe reemplazar el análisis detallado de los estudios ni la construcción de una búsqueda exhaustiva cuando el objetivo es investigación formal.

Tabla 4 Ventajas y limitaciones de Consensus para la síntesis rápida de evidencia científica

Ventajas	Limitaciones
Respuestas basadas exclusivamente en literatura científica.	No sustituye una búsqueda exhaustiva y sistemática.
Enfoque en consenso de la evidencia disponible.	Puede simplificar en exceso temas complejos o heterogéneos.
Útil para responder preguntas clínicas directas.	Menor profundidad en detalles metodológicos de cada estudio.
Proporciona enlaces a los artículos que sustentan la respuesta.	Dependiente de la cantidad y calidad de estudios disponibles.
Ahorra tiempo en la toma de decisiones preliminares.	Requiere lectura crítica posterior de los artículos originales.

Fuente: Elaboración propia con apoyo de inteligencia artificial

E. SEMANTIC SCHOLAR

Semantic Scholar es un motor de búsqueda académica impulsado por inteligencia artificial cuyo objetivo principal es priorizar relevancia científica, más que coincidencia literal de palabras clave. A diferencia de PubMed, Semantic Scholar incorpora análisis de citaciones, influencia de los artículos y relaciones entre trabajos, lo que facilita identificar estudios fundamentales dentro de un campo.

En investigación clínica, su mayor valor reside en ayudar al investigador a detectar artículos “núcleo”, es decir, aquellos que han tenido mayor impacto y que estructuran el conocimiento actual sobre un tema.

Al usar esta plataforma en el caso clínico, permitirá evidenciar estudios más citados sobre

la asociación entre AOS e hiperaldosteronismo, además de revisiones clave que han influido en guías o algoritmos diagnósticos y redes de citación que conectan estudios endocrinológicos y cardiovasculares.

Es ideal para identificar artículos clave e influyentes y para ampliar la búsqueda mediante referencias y citaciones, pero no debe sustituir a PubMed cuando se requiere una búsqueda clínica sistemática.

Tabla 5 Ventajas y limitaciones de Semantic Scholar para la identificación de artículos influyentes y redes de citación

Ventajas	Limitaciones
Prioriza artículos influyentes y altamente citados.	No utiliza MeSH ni DeCS como PubMed.
Muestra relaciones entre artículos y citaciones clave.	Puede omitir literatura regional o latinoamericana.
Útil para identificar trabajos seminales.	No está orientado a búsquedas clínicas estructuradas.
Incluye métricas de impacto y relevancia.	Requiere complementar con PubMed/BVS para exhaustividad.

F. CONNECTEDPAPERS

ConnectedPapers es una herramienta de inteligencia artificial que genera mapas visuales de literatura científica, mostrando cómo se relacionan los artículos entre sí según patrones de citación y similitud temática. Su enfoque no es responder preguntas clínicas directas, sino ayudar a comprender la estructura y evolución del conocimiento dentro de un área específica.

Si aplicamos al caso clínico se podrá visualizar cómo han evolucionado las investigaciones sobre hipertensión secundaria, identificara artículos “puente” entre Endocrinología, Medicina del Sueño y Cardiología, finalmente detectara estudios relevantes que no aparecen en búsquedas tradicionales.

Es especialmente útil para revisiones narrativas, marcos teóricos e introducciones, pero no para síntesis clínica inmediata.

Tabla 6 Ventajas y limitaciones de ConnectedPapers para la visualización de relaciones entre artículos científicos

Ventajas	Limitaciones
Visualiza relaciones entre artículos de forma intuitiva.	No responde preguntas clínicas directas.
Facilita la comprensión de la evolución del tema.	No evalúa calidad metodológica.
Ayuda a identificar artículos poco visibles pero relevantes.	Menor utilidad en decisiones clínicas urgentes.
Útil para introducciones y revisiones narrativas.	Requiere partir de al menos un artículo inicial.

Fuente: Elaboración propia con apoyo de inteligencia artificial

G. RESEARCH RABBIT

Research Rabbit es una plataforma diseñada para el descubrimiento continuo de literatura científica. Permite crear colecciones de artículos y, a partir de ellas, sugiere nuevos estudios relacionados mediante algoritmos de recomendación. A diferencia de otras herramientas, Research Rabbit está pensada para el seguimiento longitudinal de un tema, más que para búsquedas puntuales.

Si aplicamos en el caso clínico nos permite mantenernos actualizados sobre nuevas publicaciones, identificar tendencias emergentes en diagnóstico o fisiopatología y descubrir estudios recientes relacionados con los ya seleccionados.

Research Rabbit es ideal para investigadores activos que desean mantenerse actualizados en un tema específico, más que para resolver preguntas clínicas urgentes.

Tabla 7 Ventajas y limitaciones de Research Rabbit para el seguimiento continuo y descubrimiento de literatura científica

Ventajas	Limitaciones
Facilita el seguimiento continuo de un tema.	No está orientada a síntesis clínica inmediata.
Sugiere artículos relacionados automáticamente.	Depende de una selección inicial adecuada.
Útil para proyectos a mediano y largo plazo.	No sustituye una búsqueda estructurada.
Favorece el descubrimiento de literatura emergente.	Menor control metodológico de la búsqueda.

Fuente: Elaboración propia con apoyo de inteligencia artificial

7. RESOLUCIÓN DE PREGUNTAS COMPLEJAS CON OPENEVIDENCE

En la práctica clínica y en la investigación, no todas las preguntas son sencillas. Muchas veces el médico o el investigador necesita responder preguntas multidimensionales, que incluyen más de una variable o combinan varios aspectos: prevalencia, exactitud diagnóstica, pronóstico y tratamiento.

Las búsquedas tradicionales pueden requerir varias horas o días para recopilar, filtrar y analizar la literatura necesaria. Aquí es donde OpenEvidence se convierte en una herramienta valiosa: combina un modelo de lenguaje entrenado con literatura biomédica de alta calidad y la capacidad de devolver respuestas sintetizadas con referencias verificables.

¿QUÉ ES OPENEVIDENCE?

Es una plataforma de inteligencia artificial especializada en ciencias de la salud que:

- Accede a literatura científica indexada (PubMed, PMC, entre otras).
- Responde en lenguaje natural, pero con citas y enlaces directos a los artículos.
- Permite refinar la búsqueda con filtros (año, tipo de estudio, población, etc.).
- Integra información de guías clínicas, revisiones sistemáticas y ensayos.

Su mayor fortaleza está en responder preguntas clínicas complejas con soporte bibliográfico inmediato.

Aplicación al caso clínico paso a paso en OpenEvidence:

- **Ingreso de la pregunta completa:** podemos colocar la pregunta PICO tal cual está redactada, sin necesidad de abreviarla.
- **Revisión del resumen generado:** OpenEvidence devuelve un párrafo con la síntesis de resultados,
- **Validación de referencias:** Cada afirmación viene acompañada de un enlace directo a PubMed o PMC, lo que permite comprobar la fuente original.

Ejemplo de respuesta (simulada): En adultos con apnea obstructiva del sueño (AOS), la prevalencia de hiperaldosteronismo primario (PA) varía según la población estudiada, los criterios diagnósticos y la coexistencia de hipertensión. La literatura médica reciente muestra que la prevalencia de PA en pacientes con AOS e hipertensión oscila entre 11% y 27%.^[1-4] Un metaanálisis reciente reporta una prevalencia agrupada de PA del 27% (IC 95%: 24–29%) en pacientes con AOS, significativamente mayor que en pacientes sin AOS (OR $\approx$ 2).^[1] Otros estudios en cohortes de pacientes con AOS e hipertensión reportan prevalencias de PA entre 11% y 21%, con cifras más altas en hombres, pacientes jóvenes, con obesidad o con AOS moderada a severa.^[2-4]

1. Primary Aldosteronism and Obstructive Sleep Apnea: A Meta-Analysis of Prevalence and Metabolic Characteristics.

Zhang R, Cai X, Lin C, et al.

Sleep Medicine. 2024;114:8-14. doi:10.1016/j.sleep.2023.12.007.

2. Prevalence of Primary Aldosteronism in Patients With Concomitant Hypertension and Obstructive Sleep Apnea, Baseline Data of a Cohort.

Heizhati M, Aierken X, Gan L, et al.

Hypertension Research : Official Journal of the Japanese Society of Hypertension. 2023;46(6):1385-1394. doi:10.1038/s41440-023-01226-w.

3. The Prevalence of Obstructive Sleep Apnea in Patients With Primary Aldosteronism.

Sheu JY, Chang LY, Chen JY, et al.

The Journal of Clinical Endocrinology and Metabolism. 2024;109(10):2681-2691. doi:10.1210/clinem/dgae415.

4. Implementation of a Formal Sleep Center-Based Screening Protocol for Primary Aldosteronism in Patients With Obstructive Sleep Apnea.
Wright K, Mahmoudi M, Agrawal N, et al.
Surgery. 2023;173(1):59-64. doi:10.1016/j.surg.2022.05.041.

Tabla 8 Ventajas y limitaciones de OpenEvidence para la respuesta rápida a preguntas clínicas basadas en evidencia

Ventajas	Limitaciones
Orientada específicamente a preguntas clínicas.	No sustituye revisiones sistemáticas formales.
Proporciona respuestas rápidas con respaldo bibliográfico.	Menor profundidad metodológica por estudio.
Útil en contextos asistenciales con poco tiempo.	Dependiente de la disponibilidad de evidencia publicada.
Refuerza la toma de decisiones basada en evidencia.	Requiere lectura crítica posterior.

Fuente:Elaboración propia con apoyo de inteligencia artificial

Recomendación: Usar OpenEvidence cuando la pregunta involucra varios componentes clínicos, el tiempo es limitado y se necesita una visión de conjunto y se requiere citar referencias concretas en un protocolo o artículo.

Evitar usarlo como única fuente; siempre complementar con búsqueda manual en bases de datos y evaluación crítica de cada estudio.

8. RECOMENDACIONES PARA LA BÚSQUEDA DE ARTÍCULOS CON INTELIGENCIA ARTIFICIAL

El uso de inteligencia artificial en la búsqueda de literatura científica puede ser una ventaja competitiva para médicos, investigadores y estudiantes. Sin embargo, su integración debe hacerse de forma planificada, crítica y complementaria a las estrategias tradicionales de búsqueda en bases de datos.

Las siguientes recomendaciones tienen como objetivo maximizar la utilidad y minimizar los riesgos asociados.

8.1. COMBINAR LA IA CON BÚSQUEDA TRADICIONAL

Las herramientas de IA no reemplazan el acceso directo a bases de datos como PubMed, Embase, SciELO o BVS. La estrategia más segura es:

- Usar la IA para definir la pregunta de investigación, identificar términos MeSH/DeCS y generar estrategias preliminares.
- Ejecutar la búsqueda directamente en las bases de datos para verificar la exhaustividad de los resultados.

8.2. VERIFICAR SIEMPRE LAS REFERENCIAS

Algunas IA pueden “alucinar” con referencias: inventar artículos, autores o DOI. Recomendación práctica:

- Revisar cada referencia en PubMed o la base de datos original.
- Confirmar fecha, autores, revista y DOI.
- Guardar las referencias validadas en un gestor bibliográfico (Zotero, Mendeley, EndNote).

8.3. DOCUMENTAR LA ESTRATEGIA DE BÚSQUEDA

La reproducibilidad es clave en MBE. Cuando uses IA:

- Guarda capturas de pantalla del prompt y la respuesta.
- Registra los términos y operadores utilizados.
- Documenta filtros aplicados (fecha, idioma, tipo de estudio).

8.4. EVALUAR LA APLICABILIDAD CLÍNICA

La IA puede mostrar estudios de diferentes contextos geográficos o poblaciones. Antes de tomar decisiones clínicas o diseñar un protocolo, verificar:

- ¿La población estudiada es similar a la de mi contexto?
- ¿Las pruebas diagnósticas o intervenciones son viables en mi medio?
- ¿El nivel de atención y recursos disponibles son comparables?

8.5. RECONOCER LOS SESGOS PROPIOS DE LA IA

Las herramientas de IA reflejan las limitaciones de los datos con los que fueron entrenadas:

- Si prioriza estudios recientes, podría omitir evidencia histórica clave.
- Si el entrenamiento se basó mayormente en literatura anglosajona, podría subrepresentar investigaciones latinoamericanas o de países en desarrollo.

8.6. USAR IA DIFERENTE PARA OBJETIVOS DIFERENTES

Tabla 9 Herramientas de inteligencia artificial según el objetivo y momento de uso en la búsqueda de literatura científica

Objetivo clínico o investigativo	Herramienta de IA recomendada	Momento de uso en el proceso	Nivel de rigor esperado	Comentario clave para el lector
Obtener una visión rápida del estado del arte	Perplexity AI	Inicio de la búsqueda	Bajo–moderado	Útil para orientarse rápidamente y localizar artículos “semilla”; requiere verificación posterior.
Identificar y organizar estudios relevantes en tablas	Elicit	Búsqueda inicial–intermedia	Moderado	Facilita comparación entre estudios y priorización de lecturas; no sustituye PubMed.
Evaluar cómo se utiliza un artículo en la literatura	Scite AI	Análisis crítico	Moderado–alto	Permite identificar evidencia apoyada o controvertida; complementa la lectura crítica.
Conocer si existe consenso científico sobre una pregunta	Consensus	Exploración clínica rápida	Moderado	Ideal para confirmar respaldo general de una hipótesis; puede simplificar temas complejos.
Identificar artículos influyentes y trabajos seminales	Semantic Scholar	Exploración intermedia	Moderado	Útil para detectar estudios clave y redes de citación; no es búsqueda clínica estructurada.
Visualizar relaciones y evolución del conocimiento	Connected Papers	Revisión narrativa / marco teórico	Bajo–moderado	Excelente para comprender la evolución del tema y detectar artículos poco visibles.
Seguimiento continuo y descubrimiento de nueva literatura	Research Rabbit	Proyectos a mediano y largo plazo	Bajo–moderado	Ideal para mantenerse actualizado; depende de una buena selección inicial.
Responder rápidamente preguntas clínicas con respaldo científico	OpenEvidence	Práctica clínica / confirmación rápida	Moderado	Útil en contextos asistenciales; no reemplaza revisiones sistemáticas ni lectura crítica.
Búsqueda exhaustiva, reproducible y documentable	PubMed / BVS (método clásico)	Etapa final / investigación formal	Alto	Estándar de oro para protocolos, revisiones sistemáticas y publicaciones científicas.

Fuente: Elaboración propia con apoyo de inteligencia artificial

8.7. FOMENTAR EL PENSAMIENTO CRÍTICO

La IA es una herramienta de apoyo, no un sustituto del juicio clínico. Se debe asegurar de leer los artículos originales, evaluar la calidad metodológica de los estudios (usa guías CONSORT, PRISMA, STROBE), interpretar los resultados en el contexto del paciente o la población objetivo.

Como consejo final, la verdadera fortaleza de la inteligencia artificial no está en “buscar por ti”, sino en amplificar su capacidad de buscar, filtrar, sintetizar y aplicar la evidencia. Usada de forma estratégica y combinada con métodos tradicionales, la IA puede reducir drásticamente el tiempo de búsqueda y aumentar la precisión de los resultados, siempre que se mantenga la verificación humana y la lectura crítica.

8.8 EL RIESGO ÉTICO Y LA FALSIFICACIÓN DE LA EVIDENCIA ASISTIDA POR IA

El término técnico “alucinación” (generación de texto sin base factual) es un término Estratégico utilizado por la cita Apócrifa (del griego, apókrūphos: oculto, falso, de origen incierto), es una referencia bibliográfica completa (autor, título, revista, DOI) generada por un LLM que parece auténtica, pero que no existe o remite a un contenido irrelevante.

Esto ocurre porque el LLM no busca la referencia; la predice basándose en patrones de lenguaje. Sabe que tras una afirmación metodológica debe venir una cita en formato APA, y la inventa siguiendo ese patrón. Este proceso es invisible para el clínico que confía. Esto conlleva a negligencia clínica.

Tabla 10 Riesgos profesionales y éticos asociados al uso de referencias falsas generadas por inteligencia artificial

Riesgo Profesional	Escenario de Falsificación	Consecuencia Ética
Error Tipo I por Datos Falsos	La IA alucina una referencia que afirma una tasa de efectividad de un 90% para un tratamiento.	El clínico prescribe el tratamiento basándose en un dato inflado y falso, exponiendo al paciente al riesgo sin el beneficio esperado.
Omisión de Alerta	La IA, al resumir un estudio, omite o tergiversa los eventos adversos graves, citando una referencia apócrifa.	El clínico no considera el riesgo de daño (Error Tipo S/Daño) porque la alerta de seguridad fue falsificada por la IA.
Dilución de la Evidencia	En un Meta-análisis, la IA sugiere una referencia para un ECA crítico. Si esa referencia es apócrifa, todo el análisis se basa en un supuesto erróneo.	Se viola el principio de Integridad Científica, base de la credibilidad profesional.

Fuente: Elaboración propia con apoyo de inteligencia artificial

El Protocolo de verificación: Este es el protocolo que se beneficia de seguir para la verificación:

Verificación Binaria: Toda referencia generada por IA (Perplexity, Elicit, Scite, o un LLM general) debe someterse a una verificación “existe/no existe” utilizando el DOI o el título en Google Scholar o PubMed.

Verificación de Contenido: Si la referencia existe, el clínico debe abrir el abstract (y, idealmente, el texto completo) para verificar la congruencia. No basta con que el artículo exista; debe confirmar que la conclusión que la IA le atribuyó (dosis, outcome, población) sea la real.

Registro de Excepción: El experto debe crear el hábito de documentar en sus notas clínicas o de investigación: “Evidencia obtenida con IA, verificada manualmente con el DOI XXXXX”. Esta trazabilidad es su defensa ética y metodológica.

9. DE LA BÚSQUEDA PUNTUAL A LA MEDICINA BASADA EN LA EVIDENCIA DINÁMICA

El capítulo, al centrarse en una pregunta PICO y la búsqueda en bases de datos, representa el modelo tradicional y estático de la MBE: el clínico realiza una búsqueda en un momento dado para responder una pregunta.

El Desafío Estratégico: La evidencia clínica de alto impacto (ensayos de fase III) se publica a un ritmo acelerado. Una revisión sistemática (RS) está metodológicamente obsoleta en un promedio de 5.5 años, pero las guías de práctica clínica pueden tardar 18 meses en incorporar esa nueva información. Un experto no puede permitirse esta brecha.

La inteligencia artificial no solo debe agilizar la búsqueda; debe transformar el flujo de trabajo completo de la Revisión Sistemática (RS), permitiendo la Revisión Sistemática Viva (Living Systematic Review - LSR).

a. Las Revisiones Sistemáticas Vivas (LSR): Una LSR es una revisión sistemática en la que los autores se comprometen a monitorear continuamente la literatura relevante y a incorporar nuevos hallazgos en cuanto estén disponibles (generalmente cada 3 a 12 meses). Su implicación Clínica garantiza que las recomendaciones y las estimaciones del efecto (el Meta-análisis) reflejen la evidencia más actual, reduciendo el retraso en la implementación de guías clínicas.

b. La IA como Facilitador del Flujo de Trabajo (Workflow): El verdadero costo de una RS no está en la búsqueda, sino en el filtrado y la extracción de datos. La IA interviene en las siguientes etapas críticas para hacer viable una LSR:

Tabla 11 Rol de la inteligencia artificial en las distintas etapas del workflow de una revisión sistemática

Etapa del Workflow (RS)	Rol Crítico de la IA	Herramientas Típicas
Monitoreo y Detección	Alertas Programadas y Búsqueda Automática: La IA mantiene las ecuaciones de búsqueda activas en múltiples bases de datos y envía alertas sobre nuevos papers que coinciden con la estrategia avanzada.	Servicios de alertas de PubMed, Scopus, Google Scholar.
Screening (Filtrado)	Clasificación por Machine Learning: Después de eliminar duplicados, la IA lee los títulos y resúmenes y les asigna un puntaje de relevancia (ej. "probabilidad de ser un ECA relevante"). Esto permite al equipo humano priorizar y descartar rápidamente miles de artículos irrelevantes.	Rayyan, Covidence (con funciones IA): Estas plataformas están diseñadas para la gestión colaborativa del screening.
Extracción de Datos	Procesamiento de Lenguaje Natural (PNL): La IA identifica y extrae semi-automáticamente los elementos PICO	Elicit, Scite AI, OpenEvidence: Estas herramientas son fundamentales aquí, ya que estructuran la respuesta a la pregunta PICO y extraen datos específicos.
Evaluación del Sesgo	Detección de Patrones de Reporte: La IA puede ayudar a identificar inconsistencias o faltantes en el reporte (ej. falta de descripción de la secuencia de aleatorización), lo que es un indicio de un posible Riesgo de Sesgo (RoB).	Herramientas experimentales de Natural Language Processing (NLP) sobre textos completos.

Fuente: Elaboración propia con apoyo de inteligencia artificial

La IA convierte al clínico de un simple buscador de evidencia en un gestor dinámico de la información. El valor no es la velocidad de la respuesta inicial, sino la garantía de que la evidencia que utilizas hoy es la más reciente y rigurosa.

11. PREGUNTAS DE AUTOEVALUACIÓN

PREGUNTAS DE OPCIÓN MÚLTIPLE

- ¿Cuál es el principal beneficio del uso de modelos de lenguaje grande (LLM) en la búsqueda de literatura científica?
 - Reemplazar las bases de datos biomédicas tradicionales
 - Automatizar la lectura crítica de los artículos
 - Acelerar la identificación y organización inicial de evidencia relevante
 - Garantizar que la evidencia encontrada sea de alta calidad
- ¿Qué característica distingue a herramientas como Perplexity o Consensus de una búsqueda clásica en PubMed?
 - Utilizan exclusivamente términos MeSH
 - Proporcionan respuestas narrativas basadas en múltiples fuentes
 - Permiten búsquedas exhaustivas y reproducibles
 - Evalúan directamente la calidad metodológica de los estudios
- ¿Cuál de las siguientes herramientas es especialmente útil para analizar cómo un artículo es citado en la literatura científica (apoyo, contraste o mención)?
 - Elicit
 - Perplexity
 - Scite AI
 - Research Rabbit
- ¿En qué escenario es más apropiado utilizar herramientas como OpenEvidence o Consensus?
 - Elaboración de una revisión sistemática
 - Construcción de una estrategia MeSH exhaustiva
 - Confirmación rápida de si existe evidencia que respalde una hipótesis clínica
 - Análisis estadístico de los resultados de un estudio
- Según el enfoque del capítulo, ¿cuál es la relación correcta entre inteligencia artificial y Medicina Basada en la Evidencia?
 - La IA sustituye al método científico tradicional
 - La IA elimina la necesidad de lectura crítica
 - La IA complementa el método científico, mejorando la eficiencia
 - La IA garantiza conclusiones clínicas correctas

PREGUNTAS ABIERTAS

6. Explique por qué las herramientas de inteligencia artificial no deben utilizarse como único método de búsqueda bibliográfica en investigación clínica.
7. Describa dos situaciones clínicas en las que el uso de LLM puede ser especialmente útil y dos en las que no lo sería.
8. Compare brevemente el rol de Elicit y Scite AI dentro del proceso de búsqueda y análisis de literatura científica.
9. ¿Qué se entiende por “curaduría inicial de evidencia” y cuál es el papel del médico en esta etapa cuando se utilizan LLM?
10. Explique cómo puede integrarse de forma responsable la inteligencia artificial en la toma de decisiones clínicas basadas en evidencia.

RESPUESTAS

- 1: c) Los LLM permiten acelerar la identificación y organización inicial de evidencia relevante, pero no sustituyen la búsqueda clásica ni la lectura crítica.
- 2: b) Estas herramientas ofrecen respuestas narrativas basadas en múltiples fuentes, lo que las diferencia de la búsqueda estructurada clásica.
- 3: c) Scite AI se centra en analizar el contexto de las citas (apoyo, contraste o mención).
- 4: c) OpenEvidence y Consensus son útiles para confirmar rápidamente si existe respaldo científico para una hipótesis clínica.
- 5: c) La IA complementa el método científico, mejorando la eficiencia, pero no garantiza calidad ni reemplaza el juicio clínico.
6. Debe mencionarse que la IA puede omitir artículos, cometer errores bibliográficos o simplificar evidencia compleja, por lo que siempre se requiere verificación en bases de datos originales.
7. Ejemplos útiles: guardias clínicas, toma de decisiones preliminares, orientación inicial de un tema. Ejemplos no útiles: revisiones sistemáticas formales, elaboración de guías clínicas.
8. Elicit organiza estudios y facilita comparación; Scite AI evalúa cómo esos estudios son utilizados en la literatura.
9. Curaduría inicial implica seleccionar y resumir evidencia preliminar; el médico debe verificar fuentes, evaluar calidad y decidir aplicabilidad clínica.
10. Integración responsable implica usar IA como apoyo, verificar fuentes originales, realizar lectura crítica y contextualizar la evidencia al paciente real.

REFERENCIAS

- Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, Ashrafian H, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: The CONSORT-AI extension. *Lancet Digit Health*. 2020;2(10):e537–e548. Available at: [https://doi.org/10.1016/S2589-7500\(20\)30218-1](https://doi.org/10.1016/S2589-7500(20)30218-1).
- Bolaños F, Salatino A, Osborne F, Motta E. Artificial intelligence for literature reviews: Opportunities and challenges. *arXiv [Preprint]*. 2024 Aug 6. Available at: <https://arxiv.org/abs/2402.08565>.
- Marshall IJ, Wallace BC. Toward systematic review automation: A practical guide to using machine learning tools in research synthesis. *Syst Rev*. 2019;8:163. Available at: <https://doi.org/10.1186/s13643-019-1089-6>.
- Nori H, King N, McKinney SM, Carignan D, Horvitz E. Capabilities of GPT-4 on medical challenge problems. *arXiv [Preprint]*. 2023 Apr 12. Available at: <https://arxiv.org/abs/2303.13375>.
- Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, et al. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *NPJ Digit Med*. 2023;6:86. Available at: <https://doi.org/10.1038/s41746-023-00800-7>.
- Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: Can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FACCT '21)*. New York: Association for Computing Machinery; 2021. p. 610–623. Available at: <https://doi.org/10.1145/3442188.3445922>.
- Vollmer S, Mateen BA, Bohner G, Király FJ, Ghani R, Jonsson P, et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ*. 2020;368:l6927. Available at: <https://doi.org/10.1136/bmj.l6927>.
- van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CLH. ChatGPT: Five priorities for research. *Nature*. 2023;614(7947):224–226. Available at: <https://doi.org/10.1038/d41586-023-00288-7>.
- Perplexity AI. About Perplexity AI [Internet]. 2024 [cited 2025]. Available at: <https://www.perplexity.ai>.
- Elicit. Elicit: The AI research assistant [Internet]. 2024 [cited 2025]. Available at: <https://elicit.org>.
- Scite AI. How Scite works [Internet]. 2024 [cited 2025]. Available at: <https://scite.ai>.
- Consensus. Scientific consensus search engine [Internet]. 2024 [cited 2025]. Available at: <https://consensus.app>.
- Semantic Scholar. About Semantic Scholar [Internet]. 2024 [cited 2025]. Available at: <https://www.semanticscholar.org>.
- OpenEvidence. Evidence-based clinical AI search platform [Internet]. 2024 [cited 2025]. Available at: <https://openevidence.com>.