



MANUAL DE BÚSQUEDA SISTEMATIZADA Y LECTURA CRÍTICA DE EVIDENCIA EN MEDICINA, ASISTIDA POR INTELIGENCIA ARTIFICIAL



Agrarismo No. 227, Segundo Piso
Col. Escandón II Secc., D.T. Miguel Hidalgo,
C.P. 11800, Ciudad de México.

Publicado por la Dirección de Desarrollo e Integración de Medicina Basada en la Evidencia de la Dirección General de Modernización del Sector Salud.

© Copyright **Secretaría de Salud** "Derechos Reservados". Ley Federal del Derecho de Autor

Editor General

Dirección General de Modernización del Sector Salud

Dirección de Desarrollo e Integración de Medicina Basada en la Evidencia

Este manual tiene carácter exclusivamente **educativo e informativo** y está diseñado para proporcionar **orientación metodológica** sobre la búsqueda sistematizada de evidencia médica, asistida por Inteligencia Artificial (IA). **NO** constituye asesoramiento médico, diagnóstico, tratamiento, ni recomendación terapéutica específica para ningún paciente o condición clínica particular.

La **IA** es una **herramienta auxiliar** que debe ser utilizada únicamente como apoyo en el proceso de búsqueda y síntesis inicial de información científica. Las herramientas de IA presentan limitaciones inherentes que incluyen, pero no se limitan a: generación de información inexacta o ficticia ("alucinaciones algorítmicas"), sesgos inherentes en los datos de entrenamiento, limitaciones temporales en el conocimiento actualizada, incapacidad para realizar evaluación crítica metodológica completa, ausencia de comprensión del contexto clínico específico.

De conformidad con la Ley General de Salud (Artículos 32 y 33 sobre el ejercicio profesional y la responsabilidad de los profesionales de la salud) y el Reglamento de la Ley General de Salud en Materia de Prestación de Servicios de Atención Médica:

La responsabilidad de las decisiones clínicas recae **EXCLUSIVAMENTE** en el profesional médico tratante, quien debe:

1. Mantener competencia profesional actualizada en la evaluación crítica de evidencia científica.
2. Ejercer su juicio clínico independiente basado en su formación, experiencia y conocimiento médico.
3. Validar toda información obtenida mediante herramientas de IA antes de su aplicación clínica.
4. Considerar las características individuales de cada paciente (comorbilidades, preferencias, valores, contexto sociocultural).
5. Obtener consentimiento informado apropiado cuando corresponda.
6. Mantener actualizada su cédula profesional y cumplir con la normatividad vigente.

Este deslinde se fundamenta en la normatividad mexicana vigente:

- Ley General de Salud - Artículos 32 (ejercicio profesional), 33 (responsabilidad profesional), y 51 (atención médica).
- Reglamento de la Ley General de Salud en Materia de Prestación de Servicios de Atención Médica - Artículos 8, 9, y 48.
- NOM-004-SSA3-2012 "Del expediente clínico".
- Código Civil Federal - Artículos 1910-1932 sobre responsabilidad civil por hechos propios.
- Código Penal Federal - Artículos 228-230 sobre ejercicio indebido del servicio público (aplicable a médicos en instituciones públicas).

EL USO DE ESTE MANUAL IMPLICA LA ACEPTACIÓN QUE LA IA ES UNA HERRAMIENTA, NO UN SUSTITUTO DEL CONOCIMIENTO, EXPERIENCIA, Y RESPONSABILIDAD MÉDICA PROFESIONAL. La medicina sigue siendo, y permanecerá, un acto humano fundamentado en la relación médico-paciente, el conocimiento científico, la experiencia clínica, y el compromiso ético con el bienestar del paciente.

AUTORÍA INSTITUCIONAL

SECRETARÍA DE SALUD

Dr. David Kershenobich Stalnikowitz

**SUBSECRETARÍA DE INTEGRACIÓN SECTORIAL Y COORDINACIÓN DE SERVICIOS
DE ATENCIÓN MÉDICA**

Lic. Eduardo Clark García Dobarganes

DIRECCIÓN GENERAL DE MODERNIZACIÓN DEL SECTOR SALUD

Dr. José Enrique Pérez Olguín

GRUPO DE DESARROLLO

Dr. Miguel Patiño Gonzalez

Médico Especialista A

Dirección de Desarrollo e Integración de Medicina Basada en la Evidencia

Dr. José Luis Aguilar Ponce

Subdirector de Área

Servicios de Salud del Instituto Mexicano del Seguro Social para el Bienestar

Lic. José Alejandro Martínez Ochoa

Departamento de Logística de Medicina Basada en la Evidencia

Dirección de Desarrollo e Integración de Medicina Basada en la Evidencia

Dra. Pamela Stephany Guiovanna Del Moral Villavicencio

Dirección de Desarrollo e Integración de Medicina Basada en la Evidencia



ÍNDICE.

AUTORÍA INSTITUCIONAL.....	3
INTRODUCCIÓN.....	5
1. IDENTIFICACIÓN DE UNA SITUACIÓN CLÍNICA.....	8
2. ELABORACIÓN DE UNA PREGUNTA DE PRIMER PLANO.....	10
2.1 Modelo PICO: Estructuración de preguntas clínicas.....	10
3. IDENTIFICACIÓN DEL TIPO DE ESTUDIO APROPIADO PARA CONTESTAR UNA PREGUNTA PICO.....	12
3.1 Jerarquía de la evidencia científica.....	12
4. BÚSQUEDA SISTEMATIZADA DE EVIDENCIA EN MEDICINA, CON ASISTENCIA DE INTELIGENCIA ARTIFICIAL.....	15
4.1 Definición de Búsqueda Sistemática.....	15
4.2 Consideraciones Cruciales para el Entrenamiento y Validación de Modelos de IA en la Búsqueda Sistemática.....	15
4.3 Tesauros en Ciencias de la Salud: Los "Traductores" de la Evidencia.....	16
4.4 Operadores Booleanos: Los conectores que entrelazan la lógica de la búsqueda.....	17
4.5 Algoritmo de Búsqueda: Similar a una ecuación matemática.....	17
4.6 Sesgos comunes en algoritmos de búsqueda generados por IA y estrategias de mitigación.....	18
4.7 Motores de búsqueda y Metabuscadores.....	19
4.8 Sinergia Humano – IA. Enfoque <i>Researcher-in-the-Loop</i> en la revisión de Literatura.....	25
5. LECTURA CRÍTICA DE LA EVIDENCIA CIENTÍFICA.....	27
5.1 Validez interna: Metodología y herramientas de evaluación.....	27
5.2 Impacto de los resultados: Más allá del valor de *p*.....	28
5.3 Validez externa: Aplicabilidad en contextos reales.....	28
6. INTELIGENCIA ARTIFICIAL EN MEDICINA BASADA EN EVIDENCIA: LIMITACIONES Y DESAFÍOS ÉTICOS.....	34
6.1 Limitaciones técnicas de la IA.....	34
6.2 Desafíos éticos.....	34
6.3 Recomendaciones para un uso ético:.....	35
7. RECURSOS TECNOLÓGICOS EMERGENTES EN MBE.....	36
7.1 Large Language Models (LLMs) Especializados.....	36
7.2 Repositorios de Real-World Data (RWD).....	36
8. CONCLUSIONES.....	37
9. ANEXOS.....	39
9.1 Uso de la Interfaz de Chat con IA.....	39
9.2 Flujogramas.....	40
9.3 Glosario.....	42
10. BIBLIOGRAFÍA.....	45

INTRODUCCIÓN.

El presente manual, tiene como objetivo proporcionar herramientas metodológicas para la realización de búsquedas sistematizadas de literatura científica y la aplicación de una lectura crítica rigurosa a los estudios seleccionados. Cabe destacar que este documento constituye una guía de apoyo y no sustituye el conocimiento especializado, necesario para elaborar una búsqueda sistemática formal. Su finalidad es facilitar el proceso y ofrecer orientación práctica, reconociendo que un abordaje integral requiere formación específica y experiencia metodológica.

La amplia disponibilidad de información biomédica en diversas plataformas representa tanto una oportunidad como un desafío. Dada la variabilidad en la calidad y relevancia de la evidencia publicada, la identificación de estudios sólidos y pertinentes exige un proceso estructurado, reproducible y documentado. En este contexto, radica la distinción fundamental entre una búsqueda sistematizada y una búsqueda narrativa o informal: mientras ésta última puede ser útil para exploraciones iniciales, solamente un enfoque sistematizado garantiza exhaustividad y minimiza sesgos en la identificación de literatura relevante.

Este manual, incorpora el apoyo de Inteligencia Artificial (IA) para la elaboración de algoritmos de búsqueda y la aplicación de herramientas de validación interna, acorde al tipo de estudio requerido para responder preguntas clínicas. La IA optimiza la formulación de estrategias de búsqueda, mejora la identificación de evidencia pertinente y sistematiza el análisis crítico de documentos, tal como se detallará en secciones posteriores.

El proceso comienza con la selección de una situación clínica, seguida de la elaboración de una pregunta clínica o de investigación bien estructurada. Para esto, se recomienda utilizar el formato PICO (Paciente/Problema, Intervención, Comparador, Resultado); que facilita delimitar “el foco” de la búsqueda. Posteriormente se identifican términos controlados (ej.: *Medical Subject Headings* [MeSH] en PubMed o *Descriptores en Ciencias de la Salud* [DeCS] en LILACS) combinados con términos libres, operadores booleanos, truncamientos y filtros específicos por base de datos. Cada base de datos tiene características propias que se deben considerar al adaptar la estrategia de búsqueda. Algunas de las más utilizadas en salud son: PubMed/MEDLINE, Embase, Cochrane Library, Epistemonikos y LILACS. Es clave que se registre cada paso metodológico durante la búsqueda, asegurando que el proceso sea transparente y reproducible.

Una vez identificados los estudios potencialmente relevantes, se procede a evaluar tres aspectos fundamentales:

- Validez interna: Calidad metodológica del estudio.

- Consistencia y precisión de los resultados.
- Aplicabilidad al contexto clínico o poblacional.

Dado que resulta fundamental evaluar la precisión diagnóstica de los algoritmos de IA que se emplean, comparándolos con el estándar de oro, que es el análisis realizado por investigadores humanos. Se pueden incluir métricas adicionales clave, entre las que se incluyen:

- Sensibilidad: La capacidad del algoritmo para identificar correctamente todos los estudios relevantes.
- Especificidad: La capacidad del algoritmo para identificar correctamente los estudios irrelevantes.
- Precisión (Valor Predictivo): La proporción de artículos identificados como relevantes por la IA, que son realmente relevantes.
- Valor Predictivo Negativo: La proporción de artículos identificados como irrelevantes por la IA, que son realmente irrelevantes.
- Razón de Verosimilitud Positiva y Negativa: Indican cuánto aumenta o disminuye la probabilidad de que un estudio sea relevante, dado un resultado positivo o negativo del algoritmo, respectivamente.
- Razón de Odds Diagnóstica: Una medida global del rendimiento del algoritmo, que combina sensibilidad y especificidad.
- F-measure y Exactitud: Otras medidas combinadas que reflejan el equilibrio entre precisión y exhaustividad.
- Área Bajo la Curva: Representación gráfica del rendimiento del algoritmo en diferentes umbrales, siendo un indicador robusto de su capacidad discriminatoria.

Para ello, se utilizan herramientas validadas según el tipo de estudio:

- Ensayos clínicos aleatorizados: Herramienta como *RoB 2.0* de Cochrane.
- Estudios observacionales: *ROBINS-I* o listas de verificación como *STROBE*.
- Revisiones sistemáticas: *AMSTAR 2*.

Si bien se mencionan medidas de efecto (ej.: riesgo relativo [RR], *odds ratio* [OR], *hazard ratio* [HR]), este manual no profundiza en detalles técnicos. En su lugar, se proporcionan referencias clave para profundizar en estos conceptos.

Beneficios posibles de la metodología propuesta:

- Toma de decisiones informadas: Fundamentación de acciones clínicas y administrativas.



- Priorización de intervenciones: Basada en evidencia científica robusta.
- Estándares éticos y científicos: Mantenimiento de rigurosidad en entornos con variabilidad en la calidad de la información.

Este manual integra lineamientos prácticos, ejemplos aplicados y herramientas seleccionadas para guiar al lector, desde la formulación de la pregunta hasta la identificación de evidencia relevante. Con un enfoque riguroso y accesible, busca fortalecer las competencias profesionales en ámbitos clínicos, académicos e institucionales, promoviendo la autonomía y confianza en la aplicación del conocimiento.

Además de la precisión, este manual abordará cómo la IA contribuye a la reducción de la carga de trabajo y el tiempo invertido en la revisión de literatura; un aspecto crucial dada la creciente producción científica y la naturaleza intensiva en mano de obra de las revisiones sistemáticas.

1. IDENTIFICACIÓN DE UNA SITUACIÓN CLÍNICA.

La identificación de una situación clínica constituye un paso fundamental en la medicina basada en evidencia (MBE), cuyo enfoque varía según el ámbito de aplicación: académico, investigación, práctica profesional o institucional. En cada caso, se deberá adaptar la búsqueda de evidencia para garantizar su relevancia y aplicabilidad a las necesidades específicas.

Ámbitos de aplicación.

- **Ámbito académico.**
 - **Objetivo:** Formación de estudiantes y profesionales.
 - **Metodología:**
 - Uso de casos clínicos (ej.: diagnóstico diferencial de dolor torácico) para enseñar habilidades clave.
 - Priorización de escenarios comunes o complejos que integren teoría y práctica.
 - Enfoque en la formulación de preguntas estructuradas, mediante el formato PICO (*Paciente, Intervención, Comparación, Resultado*), detallado en secciones posteriores.
- **Investigación clínica.**
 - **Objetivo:** Generación de nuevo conocimiento (ej.: evaluación de la eficacia de inmunoterapia en cáncer metastásico).
 - **Metodología:**
 - Definición precisa de poblaciones de estudio, variables relevantes y metodologías robustas.
 - Apoyo en revisiones sistemáticas y ensayos controlados, para validar hipótesis o intervenciones innovadoras.
- **Práctica profesional.**
 - **Objetivo:** Atención individualizada, basada en necesidades clínicas concretas (ej.: manejo de depresión resistente en atención primaria).
 - **Metodología:**
 - Integración de tres pilares:
 - Mejor evidencia disponible.
 - Experiencia clínica del profesional.
 - Preferencias y valores del paciente.
 - Uso de herramientas ágiles para facilitar decisiones informadas, en entornos con limitaciones de tiempo.
- **Contexto institucional**
 - **Objetivo:** Mejora de la calidad asistencial (ej.: reducción de tiempos de atención en infarto agudo de miocardio).
 - **Metodología:**

- Integración de evidencia científica con datos locales, recursos disponibles y estándares de acreditación.
- Diseño de políticas clínicas efectivas y costo-eficientes.

Independientemente del contexto, es esencial evaluar:

- Población: Edad, género, comorbilidades y características individuales del paciente.
- Contexto asistencial: Nivel de atención (hospitalaria, primaria o privada).
- Objetivo clínico: Resolución de dudas diagnósticas, optimización de tratamientos o prevención de complicaciones.

Ejemplos aplicados en cirugía pediátrica.

Para ilustrar estos conceptos, se presentan las siguientes situaciones clínicas:

- Atresia de esófago: ¿Cuál es la técnica quirúrgica óptima?
- Apendicitis aguda no complicada en niños: ¿Manejo conservador vs. quirúrgico?
- Malrotación intestinal asintomática: ¿Intervención preventiva o seguimiento?
- Onfalocele y gastrosquisis: ¿Cierre primario o secuencial?

La identificación precisa de la situación clínica, es el punto de partida para aplicar la MBE de manera efectiva. Para lograrlo se sugiere:

- Definir claramente:
 - ¿A quién se atiende? (características demográficas y clínicas).
 - ¿Dónde se actúa? (recursos y nivel de atención disponibles).
 - ¿Qué se busca resolver? (diagnóstico, tratamiento o prevención).
- Traducir la evidencia en acciones concretas, adaptadas al contexto operativo.

2. ELABORACIÓN DE UNA PREGUNTA DE PRIMER PLANO.

Las preguntas clínicas pueden clasificarse en dos categorías principales según su enfoque y propósito:

Tipos de preguntas clínicas.

- **Preguntas fundamentales:**
 - Definición: Son generales y abordan conocimientos básicos (ej.: "*¿Qué es la apendicitis?*").
 - Objetivo: Obtener evidencia actualizada sobre conceptos generales de un problema de salud, sin enfoque en decisiones inmediatas.
- **Preguntas de primer plano:**
 - Definición: Surgen durante la atención clínica para resolver incertidumbres específicas (diagnósticas, terapéuticas, pronósticas, etc.).
 - Características:
 - Son concretas y se centran en aspectos particulares de una situación clínica.
 - Abordan temas donde el conocimiento aún está en evolución.

2.1 Modelo PICO: Estructuración de preguntas clínicas.

El modelo **PICO** (*Paciente/Problema, Intervención, Comparación, Resultado*) permite formular preguntas de primer plano de manera clara y sistemática, facilitando la búsqueda de evidencia aplicable. A continuación, se detalla cada componente:

P (Paciente/Población)

- Características demográficas: edad, género, etnia.
- Condición clínica: Diagnóstico específico (ej.: *cáncer de mama HER2+*), comorbilidades, gravedad.
- Contexto: Ámbito de atención (urgencias, hospitalización), ubicación geográfica (si es relevante).
- Criterios de inclusión/exclusión: (ej.: *pacientes con contraindicaciones para anticoagulantes*).

I (Intervención)

- Tipo: Fármaco, cirugía, terapia física.
- Detalles técnicos: Dosis (ej.: *5 mg/kg/día*), técnica quirúrgica (ej.: *laparoscópica*), frecuencia.
- Responsable: Especialista involucrado (ej.: *oncólogo, cirujano vascular*).

C (Comparación)

- Alternativa estándar: Tratamiento habitual, placebo o ninguna intervención.

- Diferencias técnicas: (Ej.: *cirugía abierta vs. Laparoscópica*).
- Contexto comparativo: Misma población o entorno clínico.

O (Resultado)

- Primarios: Medidas críticas como supervivencia, calidad de vida o tasa de complicaciones.
- Secundarios: Cambios en parámetros de laboratorio (ej.: *hemoglobina glicosilada*) o escalas clínicas (ej.: *escala visual analógica del dolor*).

Ejemplo aplicado:

- Situación clínica: Atresia de esófago: Técnica quirúrgica óptima.
- Pregunta PICO:
 - P: Recién nacidos con atresia de esófago.
 - I: Reparación quirúrgica mediante abordaje toracoscópico.
 - C: Reparación quirúrgica mediante toracotomía.
 - O: Seguridad (complicaciones intraoperatorias) y eficacia (tiempo de recuperación).

"En recién nacidos con atresia de esófago, ¿el abordaje toracoscópico mínimamente invasivo, comparado con la toracotomía, ofrece mayor seguridad y eficacia?"

Beneficios del modelo PICO.

- Organización de la búsqueda: Transforma incertidumbres complejas en interrogantes accionables.
- Optimización de recursos: Reduce tiempo y mejora la precisión en la identificación de evidencia.
- Integración de evidencia: Facilita la toma de decisiones informadas en ámbitos clínicos, quirúrgicos o de investigación.

3. IDENTIFICACIÓN DEL TIPO DE ESTUDIO APROPIADO PARA CONTESTAR UNA PREGUNTA PICO.

La metodología PICO es esencial para formular preguntas clínicas estructuradas en MBE. Sin embargo, la validez y precisión de las respuestas dependen de identificar el diseño de estudio adecuado, según la categoría de la pregunta. Cada tipo de investigación tiene un propósito específico y su selección correcta impacta directamente en la calidad de la práctica clínica (Tabla 1).

3.1 Jerarquía de la evidencia científica.

La evidencia científica se organiza en una jerarquía que prioriza los diseños con menor riesgo de sesgo y mayor validez interna (*Figura 1*):

1. Revisiones sistemáticas y metaanálisis.
2. Ensayos clínicos aleatorizados (ECA).
3. Estudios de cohorte.
4. Estudios de casos y controles.
5. Series de casos.
6. Opiniones de expertos.

Esta jerarquía debe interpretarse en función del tipo de pregunta clínica, ya que no todos los diseños son aplicables a todas las categorías.

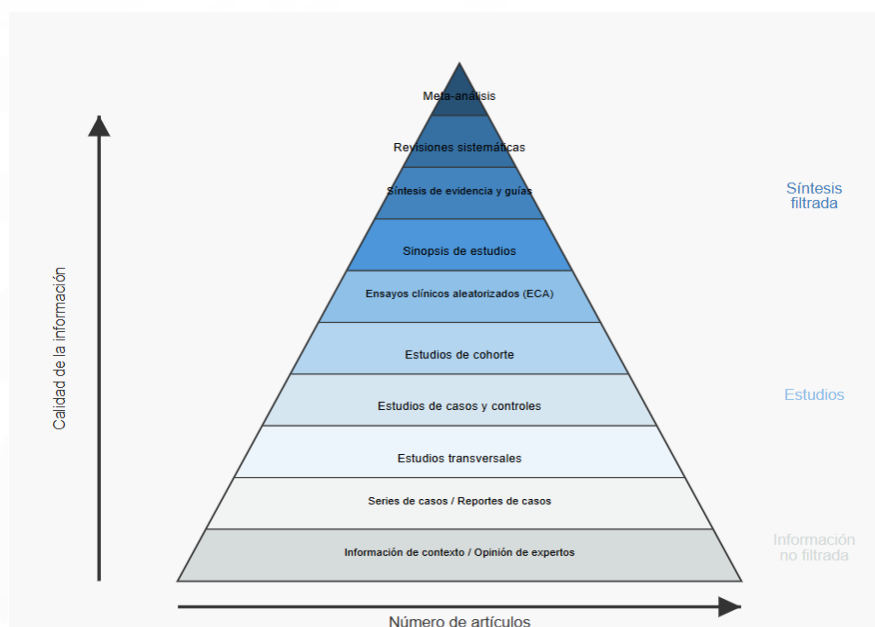


Figura 1. Jerarquía de la Evidencia Científica.

Fuente de la imagen: Generada por Claude 4 – Anthropic.

Selección del diseño de estudio según la pregunta clínica (Tabla 1).

A. Preguntas de diagnóstico.

Objetivo: Evaluar precisión, utilidad o efectividad de pruebas diagnósticas.

Diseños recomendados:

- Estudios de pruebas diagnósticas:
 - Comparan una prueba índice con un estándar de referencia ("estándar de oro").
 - Evalúan parámetros como sensibilidad y especificidad.
- Revisiones sistemáticas (RS) o metaanálisis (MET) de estudios diagnósticos:
 - Sintetizan resultados de múltiples estudios para ofrecer estimaciones cuantitativas combinadas (ej.: curvas ROC, razones de verosimilitud).

B. Preguntas de pronóstico, riesgo o etiología.

Objetivo: Explorar evolución natural de enfermedades, factores de riesgo o causas.

Diseños recomendados:

1. Estudios de cohorte:
 - Prospectivos: Siguen grupos expuestos y no expuestos a lo largo del tiempo.
 - Retrospectivos: Analizan datos históricos.
2. Estudios de casos y controles:
 - Comparan casos (con la condición) y controles (sin ella) para identificar exposiciones previas.
 - Ideales para enfermedades raras o de larga latencia.
3. RS o MET de estudios observacionales:
 - Aumentan el poder estadístico al combinar hallazgos de múltiples investigaciones.

C. Preguntas de prevención o tratamiento.

Objetivo: Evaluar eficacia, efectividad y seguridad de intervenciones.

Diseños recomendados:

1. Ensayos clínicos aleatorizados (ECA):
 - Estándar de oro: Distribución aleatoria minimiza sesgos y establece causalidad.
2. RS o MET de ECA:
 - Mejoran la precisión de las estimaciones mediante combinación de resultados.
3. Metaanálisis en red:
 - Permiten comparaciones indirectas entre intervenciones no evaluadas en un mismo estudio.

Tabla 1. Identificación del tipo de estudio apropiado para contestar una pregunta PICO.

PREGUNTA ORIENTADA A:	TIPO DE ESTUDIO
Diagnóstico	Prueba diagnóstica / Revisiones Sistemáticas (RS) o Metanálisis (MET) de pruebas diagnósticas
Pronóstico / Riesgo / Etiología	Cohorte o Casos y Controles / RS o MET de estudios observacionales
Prevención / tratamiento	Ensayo Clínico Aleatorizado (ECA) / RS o MET de ECAs, MET en Red

Fuente: Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, et al. Oxford Centre for Evidence-Based Medicine - Levels of Evidence (March 2009). Oxford: CEBM; 2009.

Importancia de la alineación PICO-diseño.

La correcta selección del tipo de estudio tiene implicaciones prácticas críticas:

- Optimización de la búsqueda bibliográfica:
 - Uso de filtros metodológicos específicos (ej.: "Randomized Controlled Trial" en PubMed).
 - Enfoque en bases de datos especializadas (ej.: Cochrane Library para ECA).
- Evaluación crítica rigurosa:
 - Aplicación de herramientas validadas (ej.: *RoB 2.0* para ECA, QUADAS-2 para estudios diagnósticos).
- Gradación de la evidencia:
 - Interpretación precisa de la calidad y fuerza de los hallazgos.
- Toma de decisiones informadas:
 - Traducción de evidencia en recomendaciones clínicas contextualizadas.

Ejemplo aplicado:

- Tema: Atresia de esófago: Técnica quirúrgica óptima.
- Pregunta PICO: En recién nacidos con atresia de esófago, ¿el abordaje toracoscópico mínimamente invasivo comparado con la toracotomía es más seguro y eficaz?
- Diseño recomendado: ECA o RS de ECA.
- Justificación: Permite evaluar causalidad y efectos directos de la intervención.

Alinear la pregunta PICO con el diseño de estudio adecuado no es un mero ejercicio teórico, sino una habilidad práctica que:

- Optimiza recursos: Reduce tiempo en búsquedas infructuosas.
- Mejora la calidad de la evidencia: Prioriza estudios con validez interna alta.
- Impacta en la atención clínica: Facilita decisiones basadas en evidencia sólida.

El diseño correcto actúa como una brújula metodológica, dirigiendo la investigación hacia respuestas confiables y aplicables.

4. BÚSQUEDA SISTEMATIZADA DE EVIDENCIA EN MEDICINA, CON ASISTENCIA DE INTELIGENCIA ARTIFICIAL.

4.1 Definición de Búsqueda Sistemática.

Una búsqueda sistematizada de evidencia médica es un proceso estructurado y reproducible para identificar, seleccionar y evaluar estudios científicos relevantes que respondan a una pregunta clínica específica. A diferencia de una búsqueda simple (como una consulta rápida en Google Scholar, (tabla 2), la búsqueda sistematizada contribuye a:

- Minimizar sesgos.
- Garantizar exhaustividad.
- Facilitar la replicabilidad.

Tabla 2. Diferencias entre búsqueda sistematizada vs. búsqueda simple.

Búsqueda Sistemática	Búsqueda Simple
Usa estrategias predefinidas (PICO).	Sin estructura clara.
Combina términos controlados y lenguaje libre.	Basada en palabras clave intuitivas.
Documenta cada paso (transparencia).	No registra métodos.
Prioriza fuentes académicas (PubMed, Cochrane).	Incluye fuentes no validadas.

Fuente: Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc. 2018 Oct 1;106(4):531-541.

4.2 Consideraciones Cruciales para el Entrenamiento y Validación de Modelos de IA en la Búsqueda Sistemática.

Cuando se utiliza un algoritmo de inteligencia artificial para la revisión de literatura, es fundamental comprender cómo ha sido entrenado y validado. La calidad y representatividad de los datos utilizados en estas fases impactan directamente en el rendimiento del modelo y en la fiabilidad de los resultados de su búsqueda. Al evaluar o emplear una herramienta de inteligencia artificial, es importante considerar la siguiente información del modelo:

Información del Conjunto de Entrenamiento:

- Nombre del conjunto de datos: Identificación específica de la literatura empleada para enseñar al algoritmo a distinguir entre estudios relevantes e irrelevantes.
- Área de estudio: Si el conjunto de entrenamiento abarca medicina general, enfermedades específicas o intervenciones particulares. Esto ayuda a comprender la especialización del modelo.
- Fuente y rango de tiempo: Las bases de datos electrónicas (ej. PubMed, Embase) de donde se extrajo la literatura y el periodo temporal que abarca este conjunto de datos.
- Tipo de publicación: Si el entrenamiento se realizó con resúmenes o textos completos.

- Número de literaturas: El total de artículos en el conjunto de entrenamiento y el número de los que fueron clasificados como incluidos.
- Método de entrenamiento: Si el modelo fue entrenado de forma supervisada (con etiquetas humanas), semi-supervisada o no supervisada. Este detalle es crítico para entender la base de su aprendizaje.

Información del Conjunto de Validación:

- Nombre del conjunto de datos: Identificación de la literatura utilizada para probar el rendimiento del algoritmo una vez entrenado.
- Área, Fuente y Rango de tiempo: Similares a los del conjunto de entrenamiento, pero aplicados a los datos de validación.
- Tipo de publicación y Número de literaturas: Detalla la naturaleza y cantidad de los artículos en el conjunto de validación, así como el número de aquellos considerados elegibles.
- Estándar de oro: Describe el proceso de revisión manual, realizado por investigadores humanos, que se utilizó como referencia para evaluar el rendimiento del algoritmo.

Comprender estos aspectos es vital para asegurar que la IA sea una herramienta adecuada y fiable para la pregunta de investigación específica, minimizando sesgos derivados de un entrenamiento o validación inadecuados.

4.3 Tesauros en Ciencias de la Salud: Los "Traductores" de la Evidencia.

Los tesauros son sistemas de vocabulario controlado que estandarizan términos médicos, evitando ambigüedades. Funcionan como "traductores" que agrupan conceptos, sin importar cómo se escriban en los estudios. Los más usados son:

MeSH (Medical Subject Headings):

- Desarrollado por la Biblioteca Nacional de Medicina de EE. UU. para PubMed.
- Ejemplo: "Infarto", "Heart Attack", "IAM" → Todos se agrupan bajo "Myocardial Infarction"[MeSH].

EMTREE (Embase):

- Tesauro de Embase, base de datos europea especializada en farmacología y enfermedades raras.
- Incluye términos detallados para fármacos (ej.: "Keytruda®" se indexa como "pembrolizumab"/EMTREE).

DeCS (Descriptores en Ciencias de la Salud):

- Creado por BIREME para bases como LILACS (América Latina).

- Incluye términos en español, portugués, inglés y francés (ej.: "atención primaria de salud" es un descriptor oficial).

Ventajas del uso de tesauros en búsquedas.

Los tesauros (vocabularios controlados) son herramientas clave para optimizar la precisión y coherencia en la recuperación de evidencia:

- Evitan duplicados: Unifican términos sinónimos bajo un concepto único (ej.: "Infarto", "Heart Attack" → "Myocardial Infarction" [MeSH]).
- Mejoran la precisión: Utilizan términos estandarizados que abarcan variaciones lingüísticas.
- Adaptación contextual:
 - Emtree (Embase): Ideal para fármacos y estudios europeos.
 - DeCS (LILACS): Especializado en evidencia regional de América Latina.

4.4 Operadores Booleanos: Los conectores que entrelazan la lógica de la búsqueda.

Los operadores booleanos (AND, OR, NOT) son herramientas que te permiten combinar términos y ajustar la sensibilidad o especificidad de una búsqueda:

- OR: Amplía resultados (ej.: "Child" OR "Pediatric"). La búsqueda incluye todos los registros que tengan cualquiera de los 2 términos. Con ello buscamos sensibilidad, es decir la mayor cantidad de resultados posibles.
- AND: Restringe a coincidencias (ej. "Aspirin" AND "Myocardial Infarction"). La búsqueda incluye los registros que contengan ambos términos. Con ello buscamos especificidad, es decir, mayor precisión y pertinencia en los resultados.
- NOT: Excluye términos (ej.: "Diabetes" NOT "Type 2"). La búsqueda incluye uno o más términos, pero excluye otro(s). Con ello se busca especificidad, mayor precisión.

Ejemplo de ecuación booleana:

("Thoracoscopy" [MeSH] OR "Minimally Invasive Surgical Procedures" [MeSH]) AND ("Postoperative Complications" [MeSH] OR "Anastomotic Leak" [MeSH])

4.5 Algoritmo de Búsqueda: Similar a una ecuación matemática.

Un algoritmo de búsqueda es una fórmula que traduce tu pregunta clínica (usualmente estructurada en PICO) en una cadena de términos y operadores. Cada componente se relaciona como una variable:

- Población (P): Términos demográficos o diagnósticos (ej.: "Infant, Newborn"[MeSH]).

- Intervención (I) y Comparación (C): Técnicas, fármacos o procedimientos (ej.: "Thoracoscopy"[MeSH] vs. "Thoracotomy"[MeSH]).
- Resultados (O): Métricas de evaluación (ej.: "Mortality"[MeSH]).

Estructura general:

(Población) AND (Intervención) AND (Comparación) AND (Resultados)

4.6 Sesgos comunes en algoritmos de búsqueda generados por IA y estrategias de mitigación.

A. Sesgo de selección de términos (*Language Bias*)

- Qué ocurre: Uso exclusivo de términos en un idioma o sin incluir vocabulario controlado.
- Cómo evitarlo:
 - Incluya términos MeSH/Descriptores específicos (ej.: "Esophageal Atresia" [MeSH]).
 - Combine lenguaje libre y controlado:

("Thoracoscopy" [MeSH] OR "Minimally Invasive Surgical Procedures" [MeSH] OR thoracoscop)*

B. Sesgo de publicación (*Publication Bias*)

- Qué ocurre: Exclusión de evidencia antigua o literatura gris.
- Cómo evitarlo:
 - Elimine límites temporales (salvo que sean necesarios).
 - Incluya términos para literatura gris (ej.: tesis, ClinicalTrials.gov).

C. Sesgo de búsqueda restringida (*Database Bias*)

- Qué ocurre: Búsqueda limitada a una sola base de datos.
- Cómo evitarlo:
 - Genere estrategias adaptables a múltiples plataformas (ej.: PubMed, Embase).
 - Sugiera bases complementarias (ej.: LILACS para América Latina).

D. Sesgo de diseño metodológico (*Study Design Bias*)

- Qué ocurre: No priorizar diseños robustos (ej.: ECA).
- Cómo evitarlo:
 - Aplique filtros metodológicos:

("Appendicitis" [MeSH] AND "Anti-Bacterial Agents" [MeSH]) AND ("Appendectomy" [MeSH] OR "Conservative Treatment" [MeSH])

E. Sesgo de especificidad (*Over-Specificity Bias*)

- Qué ocurre: Términos demasiado restrictivos (ej.: "neonatos prematuros").
- Cómo evitarlo:
 - Use operadores OR para ampliar conceptos:

("Infant, Newborn" [MeSH] OR "Infant, Premature" [MeSH])

F. Sesgo de actualización (*Currency Bias*)

- Qué ocurre: Omisión de guías clínicas recientes.
- Cómo evitarlo:
 - Incluya filtros como *"Practice Guideline"* [Publication Type].

4.7 Motores de búsqueda y Metabuscadores.

Un motor de búsqueda (ej.: PubMed o Embase) indexa documentos científicos bajo criterios predefinidos. Estos sistemas ofrecen acceso a literatura especializada, pero requieren que realices búsquedas en cada base por separado.

Un metabuscador (ej.: Epistemonikos o Tripdatabase) integra múltiples motores en una sola interfaz, optimizando tiempo y reduciendo sesgos de selección (Tabla 3).

Elección clave:

- Motores: Mayor control con filtros avanzados (tipo de estudio o idioma).
- Metabuscadores: Visión panorámica de la evidencia disponible.

Tabla 3. Comparativa de algunas bases de datos en medicina.

Base de datos	Características principales	Principales tipos de documentos recuperables	Link
PubMed	Acceso gratuito. Indexa MEDLINE y artículos biomédicos. Filtros por tipo de estudio.	Artículos científicos, revisiones, ensayos clínicos, guías clínicas.	https://pubmed.ncbi.nlm.nih.gov/
Embase	Enfocado en farmacología y biomedicina. Incluye resúmenes de conferencias. Requiere suscripción.	Estudios preclínicos, informes de seguridad de medicamentos, artículos europeos.	https://www.embase.com/
SciELO	Acceso abierto. Especializado en revistas de América Latina y el	Artículos científicos, revisiones sistemáticas, tesis (en español/inglés/portugués).	https://scielo.org/es/



Base de datos	Características principales	Principales tipos de documentos recuperables	Link
	Caribe. Multilingüe.		
LILACS	Literatura científica en salud de América Latina. Indexa tesis y documentos técnicos.	Artículos grises, informes gubernamentales, estudios observacionales.	https://lilacs.bvsalud.org/es/
Google Scholar	Amplio alcance (no solo científico). Incluye citaciones y patentes. Sin filtros avanzados.	Artículos, libros, tesis, documentos no revisados por pares.	https://scholar.google.com/
Cochrane	Enfocado en revisiones sistemáticas y ensayos clínicos. Herramientas para evaluación crítica.	Revisiones Cochrane, protocolos, resúmenes de evidencia (CENTRAL).	https://www.cochranelibrary.com/es/
Epistemonikos	Metabuscador que integra múltiples fuentes (Cochrane, PubMed). Clasifica por nivel de evidencia.	Revisiones sistemáticas, resúmenes estructurados.	https://www.epistemonikos.org/es
Tripdatabase	Búsqueda basada en PICO. Filtros por relevancia clínica. Acceso parcial gratuito.	Guías de práctica clínica, artículos de alto impacto, resúmenes de evidencia.	https://www.tripdatabase.com/
ClinicalTrials.gov	Registro global de ensayos clínicos. Información sobre estudios en curso o completados.	Protocolos de ensayos, resultados preliminares, datos de seguridad.	https://clinicaltrials.gov/

Fuente: Análisis comparativo elaborado por Claude 4 (Anthropic) con base en características técnicas y funcionales de cada plataforma, 2025.

Tema: Atresia de esófago: Técnica quirúrgica óptima.

- Pregunta PICO: En recién nacidos con atresia de esófago, ¿el abordaje toracoscópico mínimamente invasivo comparado con la toracotomía es más seguro y eficaz?
- Estudio para responder la pregunta: RS o MET.
 - Instrucciones dadas al modelo de IA para generar el algoritmo.

Investigador: "Elabora dos algoritmos de búsqueda bibliográfica basados en la siguiente pregunta PICO:

- Población: Recién nacidos con atresia de esófago.
- Intervención: Abordaje toracoscópico mínimamente invasivo.
- Comparador: Toracotomía.
- Desenlaces: Seguridad y eficacia.

Requisitos específicos:

- Desarrolla un algoritmo para PubMed y otro para Embase.
- Estructura cada algoritmo utilizando operadores booleanos (AND, OR, NOT) y sintaxis específica de cada base de datos.
- Incluye términos MeSH/Emtree y palabras clave en texto libre para maximizar la sensibilidad.
- Incorpora filtros metodológicos validados para identificar revisiones sistemáticas y metaanálisis.
- Aplica estrategias para minimizar sesgos de selección y publicación (incluir literatura gris, no restringir por idioma o fecha).
- Prioriza la sensibilidad sobre la especificidad para garantizar una búsqueda exhaustiva.
- Documenta el número de resultados obtenidos con cada algoritmo.
- Sugiere términos adicionales para refinar la búsqueda si fuera necesario."

Respuesta del modelo de IA: Algoritmos de búsqueda bibliográfica para Atresia de Esófago para Embase y Pubmed.

Algoritmo para Embase.

('esophagus atresia'/exp OR 'esophageal atresia':ti,ab,kw OR 'oesophageal atresia':ti,ab,kw OR 'atresia of esophagus':ti,ab,kw OR 'atresia of oesophagus':ti,ab,kw OR 'esophageal gap':ti,ab,kw OR 'esophageal defect':ti,ab,kw OR 'TEF':ti,ab,kw OR 'tracheoesophageal fistula':ti,ab,kw OR 'tracheo-esophageal fistula':ti,ab,kw OR 'TOF':ti,ab,kw) AND ('thoracoscopy'/exp OR 'minimally invasive surgery'/exp OR 'video assisted thoracoscopic surgery'/exp OR 'thoracoscop*':ti,ab,kw OR 'minimally invasive':ti,ab,kw OR 'video-assisted':ti,ab,kw OR 'endoscopic surgery':ti,ab,kw OR 'VATS':ti,ab,kw OR 'minimal access':ti,ab,kw OR 'laparoscopic approach':ti,ab,kw OR 'laparoscopic repair':ti,ab,kw OR 'endoscopic repair':ti,ab,kw) AND ('thoracotomy'/exp OR 'thoracotom*':ti,ab,kw OR 'open surgery':ti,ab,kw OR 'open approach':ti,ab,kw OR 'open repair':ti,ab,kw OR 'conventional surgery':ti,ab,kw OR 'traditional approach':ti,ab,kw OR 'surgical repair':ti,ab,kw OR 'standard repair':ti,ab,kw) AND ('meta analysis'/exp OR 'systematic review'/exp OR 'metaanaly*':ti,ab,kw OR 'meta analysis':ti,ab,kw OR 'systematic review':ti,ab,kw OR 'systematic overview':ti,ab,kw OR 'systematic literature review':ti,ab,kw OR 'cochrane review':ti,ab,kw OR 'umbrella review':ti,ab,kw OR 'meta synthesis':ti,ab,kw OR 'metasynthesis':ti,ab,kw OR 'integrative review':ti,ab,kw OR 'evidence synthesis':ti,ab,kw) AND [embase]/lim

Algoritmo para PubMed.

("Esophageal Atresia"[MeSH] OR "esophageal atresia"[tiab] OR "oesophageal atresia"[tiab] OR "atresia of esophagus"[tiab] OR "atresia of oesophagus"[tiab] OR "esophageal gap"[tiab] OR "esophageal defect"[tiab] OR "TEF"[tiab] OR "tracheoesophageal fistula"[tiab] OR "tracheo-esophageal fistula"[tiab] OR "TOF"[tiab]) AND ("Thoracoscopy"[MeSH] OR "Thoracic Surgery, Video-Assisted"[MeSH] OR thoracoscop*[tiab] OR "minimally invasive"[tiab] OR "video-assisted"[tiab] OR "endoscopic surgery"[tiab] OR VATS[tiab] OR "minimal access"[tiab] OR "laparoscopic approach"[tiab] OR "laparoscopic repair"[tiab] OR "endoscopic repair"[tiab]) AND ("Thoracotomy"[MeSH] OR thoracotom*[tiab] OR "open surgery"[tiab] OR "open approach"[tiab] OR "open repair"[tiab] OR "conventional surgery"[tiab] OR "traditional approach"[tiab] OR "surgical repair"[tiab] OR "standard repair"[tiab]) AND ("Meta-Analysis"[pt] OR "Meta-Analysis as Topic"[MeSH] OR metaanaly*[tiab] OR meta-analy*[tiab] OR "meta analysis"[tiab] OR "systematic review"[tiab] OR "systematic overview"[tiab] OR "systematic literature review"[tiab] OR "Systematic Review"[Publication Type] OR "Review Literature as Topic"[MeSH] OR "Review"[Publication Type] OR "Cochrane Database Syst Rev"[Journal] OR "umbrella review"[tiab] OR "meta-synthesis"[tiab] OR metasyntesis[tiab] OR "integrative review"[tiab])

Resultados obtenidos en PubMed: 39.

Resultados que cumplen con los criterios solicitados: 5. Como se observa en la Imagen 1.

Imagen 1.

PMID	Title	Authors	Citation	First Author	Journal/Book	Publication	Create Date	PMCID	NIHMS ID	DOI
33002943	Thoracoscop	Drevin G, An	Ann Surg. 20	Drevin G	Ann Surg	2021	01/10/2020			10.1097/SLA.00000000000004239
31359222	Thoracoscop	Way C, Wayr	Pediatr Surg	Way C	Pediatr Surg	2019	31/07/2019			10.1007/s00383-019-04527-9
23172569	Thoracoscop	Borruto FA, I	Eur J Pediatr	Borruto FA	Eur J Pediatr	2012	23/11/2012			10.1055/s-0032-1329711
27472740	Outcomes of	Yang YF, Don	Medicine (B	Yang YF	Medicine (B	2016	30/07/2016	PMC5265877		10.1097/MD.00000000000004428
28914345	Comparison	Wu Y, Kuang	Pediatr Surg	Wu Y	Pediatr Surg	2017	16/09/2017			10.1007/s00383-017-4153-9

Nota: Obtenida de PubMed solicitando salvar los documentos seleccionados del total en formato CSV.

Información adicional para la búsqueda.

- Para añadir términos de población (si es necesario aumentar especificidad): ("Infant, Newborn"[MeSH] OR newborn*[tiab] OR neonate*[tiab] OR infant*[tiab] OR baby[tiab] OR babies[tiab])
- Para añadir términos de resultados (si es necesario aumentar especificidad): (safety[tiab] OR efficacy[tiab] OR effective*[tiab] OR outcome*[tiab] OR complication*[tiab] OR mortality[tiab] OR morbidity[tiab] OR leak*[tiab] OR stenosis[tiab] OR "operative time"[tiab] OR "length of stay"[tiab] OR "hospital stay"[tiab])
- Para búsqueda en literatura gris (OpenGrey): Se sugiere realizar una búsqueda en OpenGrey, repositorio especializado en literatura gris, utilizando la siguiente estrategia: ("esophageal atresia" OR "oesophageal atresia") AND (thoracoscopy OR "thoracoscopic surgery" OR "minimally invasive") AND (thoracotomy OR "open surgery")

Para registros de ensayos clínicos (ClinicalTrials.gov): (esophageal atresia OR oesophageal atresia) AND (thoracoscopic OR thoracoscopy OR minimally invasive) AND (thoracotomy OR open surgery)

Esta estrategia prioriza la sensibilidad sobre la especificidad para garantizar una búsqueda exhaustiva y reduce al mínimo los riesgos de sesgo de selección y publicación.

Algoritmo 2 para PubMed que enfatiza la especificidad y recupera solo metanálisis.

Investigador: "Elabora un algoritmo de búsqueda bibliográfica basado en la siguiente pregunta PICO:

- Población: Recién nacidos con atresia de esófago.
- Intervención: Abordaje toracoscópico mínimamente invasivo.
- Comparador: Toracotomía.
- Desenlaces: Seguridad y eficacia.

Requisitos específicos:

- Desarrolla un algoritmo exclusivamente para PubMed.
- Estructura el algoritmo utilizando operadores booleanos (AND, OR, NOT) y sintaxis específica de PubMed.
- Utiliza términos MeSH precisos y palabras clave altamente específicas.
- Incorpora filtros metodológicos estrictos para identificar únicamente metaanálisis (no revisiones sistemáticas generales).
- Prioriza la especificidad sobre la sensibilidad para obtener resultados más precisos y reducir el ruido.
- Incluye términos específicos de población que identifiquen claramente a recién nacidos.
- Añade filtros temporales para obtener sólo la evidencia más reciente (últimos 10 años).
- Estructura el algoritmo para que pueda copiarse y pegarse directamente en el buscador de PubMed."

Respuesta del modelo de IA: Algoritmo para PubMed (Alta Especificidad - Sólo Metaanálisis).

("Esophageal Atresia"[Majr] OR ("esophageal atresia"[ti] AND (newborn*[ti] OR neonate*[ti] OR infant*[ti]))) AND ("Thoracoscopy"[Majr] OR "Thoracic Surgery, Video-Assisted"[Majr] OR thoracoscop*[ti] OR "minimally invasive"[ti]) AND ("Thoracotomy"[Majr] OR thoracotom*[ti] OR "open surgery"[ti]) AND ("Meta-Analysis"[pt] OR metaanaly*[ti] OR meta-analy*[ti]) AND ("last 10 years"[PDat])

Resultados obtenidos: 3.

Resultados que cumplen con los criterios solicitados: 3. Como se observa en la Imagen 2.

Imagen 2.

PMID	Title	Authors	Citation	First Author	Journal/Boo	Publication	Create Date	PMCID	NIHMS ID	DOI
33002943	Thoracoscop	Drevin G, An	Ann Surg. 20	Drevin G	Ann Surg	2021	01/10/2020			10.1097/SLA.0000000000004239
31359222	Thoracoscop	Way C, Wayr	Pediatr Surg	Way C	Pediatr Surg	2019	31/07/2019			10.1007/s00383-019-04527-9
27472740	Outcomes of	Yang YF, Don	Medicine (B	Yang YF	Medicine (B	2016	30/07/2016	PMC5265877		10.1097/MD.0000000000004428

Nota: Obtenida de PubMed. Solicitando salvar los documentos seleccionados del total en formato CSV.

Este algoritmo está diseñado para maximizar la especificidad mediante:

- Uso de términos MeSH como tema principal [Majr] en lugar de términos MeSH generales.
- Restricción de búsqueda a términos en el título [ti] para los conceptos principales.
- Especificación de la población pediátrica mediante la combinación con términos de recién nacidos.
- Limitación estricta a metaanálisis mediante filtros de publicación y términos en el título.
- Restricción temporal a los últimos 10 años para obtener evidencia actualizada.

Esta estrategia priorizará encontrar estudios altamente relevantes (específicos) a expensas de potencialmente perder algunos estudios menos específicos, pero posiblemente relevantes.

Ejemplo aplicado:

Tema: Atresia de esófago: Técnica quirúrgica óptima.

- Pregunta PICO: En recién nacidos con atresia de esófago, ¿el abordaje toracoscópico mínimamente invasivo comparado con la toracotomía es más seguro y eficaz?
- Estudio para responder la pregunta: RS o MET.

Algoritmo:

("Esophageal Atresia"[MeSH] OR "esophageal atresia"[tiab] OR "oesophageal atresia"[tiab] OR "atresia of esophagus"[tiab] OR "atresia of oesophagus"[tiab] OR "esophageal gap"[tiab] OR "esophageal defect"[tiab] OR "TEF"[tiab] OR "tracheoesophageal fistula"[tiab] OR "tracheo-esophageal fistula"[tiab] OR "TOF"[tiab]) AND ("Thoracoscopy"[MeSH] OR "Thoracic Surgery, Video-Assisted"[MeSH] OR thoracoscop*[tiab] OR "minimally invasive"[tiab] OR "video-assisted"[tiab] OR "endoscopic surgery"[tiab] OR VATS[tiab] OR "minimal access"[tiab] OR "laparoscopic approach"[tiab] OR "laparoscopic repair"[tiab] OR "endoscopic repair"[tiab]) AND ("Thoracotomy"[MeSH] OR thoracotom*[tiab] OR "open surgery"[tiab] OR "open approach"[tiab] OR "open repair"[tiab] OR "conventional surgery"[tiab] OR "traditional approach"[tiab] OR "surgical repair"[tiab] OR "standard repair"[tiab]) AND ("Meta-Analysis"[pt] OR "Meta-Analysis as Topic"[MeSH] OR metaanaly*[tiab] OR meta-analy*[tiab] OR "meta analysis"[tiab] OR "systematic review"[tiab] OR "systematic overview"[tiab] OR "systematic literature review"[tiab] OR "Systematic Review"[Publication Type] OR "Review Literature as Topic"[MeSH] OR "Review"[Publication Type] OR "Cochrane Database Syst Rev"[Journal] OR "umbrella review"[tiab] OR "meta-synthesis"[tiab] OR metasynthesis[tiab] OR "integrative review"[tiab])

- Resultados totales: 39.
- Resultados potencialmente útiles: 5.

La búsqueda sistematizada actúa como un puente esencial entre las preguntas clínicas y la evidencia confiable. Al emplear herramientas como tesauros (MeSH, Emtree, DeCS), operadores booleanos y diseño de algoritmos, es posible transformar incertidumbres en respuestas prácticas y aplicables. Para optimizar este proceso, es clave combinar motores de búsqueda especializados (como PubMed o Cochrane), que priorizan precisión y rigor metodológico, con metabuscadores (como Epistemonikos o TripDatabase), que amplían el alcance hacia guías clínicas y revisiones rápidas.

Es importante considerar que, aunque realizar un algoritmo específico disminuye considerablemente el número de documentos a revisar y reduce la carga de trabajo del investigador, esta estrategia puede ocasionar que no se recuperen documentos potencialmente útiles para el tema revisado. Este equilibrio entre especificidad y sensibilidad constituye un desafío fundamental en la metodología de búsqueda: mientras las estrategias altamente específicas mejoran la precisión y relevancia inmediata de los resultados, pueden omitir estudios valiosos que utilizan terminología menos convencional o que abordan el tema desde perspectivas complementarias. Por ello, en investigaciones críticas como metaanálisis, puede ser aconsejable implementar tanto estrategias de alta sensibilidad como de alta especificidad, comparando posteriormente los resultados para garantizar una cobertura óptima de la evidencia disponible.

La planificación rigurosa del tiempo es un factor esencial para garantizar la calidad en cada etapa del proceso de investigación. Cada fase requiere una dedicación meticulosa para minimizar sesgos y errores. Este enfoque sistemático se complementa con la elección estratégica de plataformas de búsqueda, adaptadas al contexto y objetivos del estudio. Por ejemplo, mientras PubMed es óptimo para acceder a estudios primarios y metaanálisis, Tripdatabase prioriza guías de práctica clínica basadas en consensos; agilizando la identificación de evidencia aplicable. Además, la inclusión de fuentes como LILACS, especializada en literatura biomédica de América Latina, enriquece la relevancia geográfica y cultural de los hallazgos. Con práctica, esta integración de gestión temporal y selección crítica de recursos no solo optimiza el proceso investigativo, sino que también fortalece la toma de decisiones clínicas éticas e informadas.

4.8 Sinergia Humano – IA. Enfoque *Researcher-in-the-Loop* en la revisión de Literatura

La asistencia de la inteligencia artificial en la búsqueda sistematizada no implica una automatización completa, sino una colaboración activa entre el algoritmo y el investigador, un enfoque conocido como ***Researcher-in-the-Loop***. Este modelo incrementa las fortalezas de ambas partes: la capacidad de procesamiento y aprendizaje de patrones de la IA, y el juicio crítico, el conocimiento contextual y la experiencia clínica del ser humano.

El proceso típicamente funciona de la siguiente manera:

1. **Entrenamiento Inicial:** Antes de iniciar la revisión, el algoritmo de IA necesita ser entrenado con un pequeño conjunto de artículos que ya han sido clasificados manualmente como relevantes o irrelevantes. Cuanto mayor sea este conocimiento previo, mejor entrenado estará el algoritmo al inicio del proceso.
2. **Clasificación y Proposición:** Una vez entrenado, la herramienta de IA utiliza algoritmos de *machine learning* (ej: Naïve Bayes con TF-IDF) para clasificar todos los artículos no etiquetados según su probabilidad de ser relevantes. El sistema entonces propone al revisor el artículo con la mayor probabilidad de ser relevante.

3. Decisión y re-entrenamiento iterativo: El revisor lee el título y el resumen del artículo propuesto y decide sobre su elegibilidad, (relevante o irrelevante) según los criterios de inclusión. Esta decisión humana es retroalimentada al algoritmo, que la incorpora como nuevo conocimiento y reajusta su clasificación, proponiendo un nuevo artículo en la siguiente iteración. Este ciclo se repite hasta alcanzar un criterio de parada predefinido.

Beneficios de la Interacción:

Eficiencia Acelerada: El algoritmo prioriza la presentación de artículos potencialmente relevantes, lo que reduce considerablemente el número total de artículos que el revisor debe evaluar manualmente. Se ha observado que, con este enfoque, se puede revisar una fracción significativa de la literatura total, ahorrando una considerable cantidad de tiempo.

Adaptación Dinámica: El aprendizaje continuo del algoritmo le permite adaptarse a la especificidad de la pregunta de investigación y a la terminología del campo.

Desafíos en la Interacción

- **Influencia de la subjetividad o errores del revisor:** Las decisiones iniciales del revisor pueden influir en el entrenamiento del algoritmo. Errores sistemáticos o sesgos personales podrían desviar la clasificación de la IA, llevando a la omisión de artículos relevantes. **Solución:** Realizar un entrenamiento del revisor previo para asegurar una comprensión uniforme de los criterios de elegibilidad. Implementar revisión doble o parcial (ej: el 20% de los artículos revisados por el primer revisor) y calcular el acuerdo inter-revisor (Kappa de Cohen) para evaluar la fiabilidad de las propuestas.
- **Elección del Criterio de Parada:** Detener la revisión de manera óptima es crucial para equilibrar el ahorro de tiempo con la exhaustividad. Un criterio de parada mal definido podría resultar en la omisión de artículos relevantes (detenerse antes de tiempo) o en una pérdida de tiempo al revisar muchos artículos irrelevantes (detenerse demasiado tarde). **Solución:** Formular un criterio de parada basado en datos predefinido (ej: un número específico de artículos irrelevantes consecutivos propuestos por la IA). Solicitudes futuras deberían explorar el efecto de diferentes cantidades de conocimiento previo y número de artículos que definen el criterio de parada.

5. LECTURA CRÍTICA DE LA EVIDENCIA CIENTÍFICA.

La lectura crítica de evidencia médica es un proceso estructurado que evalúa la solidez metodológica, relevancia clínica y aplicabilidad de los hallazgos científicos. Este análisis se centra en los siguientes tres componentes fundamentales:

5.1 Validez interna: Metodología y herramientas de evaluación.

La validez interna determina si un estudio establece una relación causal confiable entre una intervención y un resultado, minimizando errores sistemáticos (tabla 4). Su evaluación depende de:

A. Diseño metodológico.

1. Ensayos clínicos aleatorizados (ECA):
 - Estándar de oro: Control de factores de confusión mediante aleatorización y enmascaramiento.
2. Estudios observacionales:
 - Útiles para explorar asociaciones, pero con mayor riesgo de sesgos.
3. Revisiones sistemáticas y metaanálisis:
 - Priorizan transparencia en la selección y análisis de estudios.
4. Estudios de pruebas diagnósticas:
 - Evalúan precisión frente a un estándar de referencia (ej.: biopsia para cáncer).

Tabla 4. Herramientas para evaluar validez interna según tipo de estudio.

Tipo de Estudio	Herramienta de evaluación	Descripción	Link
Revisiones sistemáticas y metaanálisis	AMSTAR 2	Evalúa calidad metodológica mediante 16 ítems (p. ej., protocolo registrado, análisis de sesgos).	https://amstar.ca/
Ensayos clínicos aleatorizados (ECA)	RoB 2	Analiza riesgo de sesgo en 5 dominios (p. ej., aleatorización).	https://methods.cochrane.org/bias/resources/rob-2-revised-cochrane-risk-bias-tool-randomized-trials
Estudios observacionales	ROBINS-I	Evalúa riesgo de sesgo en estudios no aleatorizados.	https://methods.cochrane.org/robins-i
Estudios de pruebas diagnósticas	QUADAS-2	Valida calidad en 4 dominios (p. ej., selección de participantes).	https://www.bristol.ac.uk/population-health-sciences/projects/quadas/quadas-2/

Nota: Estas son solamente algunas de las herramientas más utilizadas, pero existen otras adaptadas a diseños específicos. Fuente: Análisis comparativo elaborado por Claude 4 (Anthropic) con base en características técnicas y funcionales de cada herramienta, 2025.

5.2 Impacto de los resultados: Más allá del valor de p^* .

El valor de " p " indica la probabilidad de obtener resultados observados si no hay efecto real. Sin embargo:

Limitaciones del valor de " p ".

- No mide el tamaño del efecto:
 - Un $p < 0.05$ no indica si una reducción del riesgo es clínicamente relevante (ej.: fármaco que reduce fiebre en 0.1°C).
- No confirma causalidad:
 - Solo refleja rareza estadística, no prueba que la intervención funcione.
- Evite el término "estadísticamente significativo":
 - Priorice la interpretación del impacto clínico (ej.: *número necesario a tratar [NNT]*).

Ejemplo:

Un estudio con 10,000 pacientes muestra que un fármaco reduce la mortalidad del 10% al 9% (valor de $p = 0.03$). Aunque hay diferencia estadística, salvarías 1 vida por cada 100 pacientes tratados. ¿Vale la pena el costo y los efectos secundarios? La respuesta depende del contexto clínico.

5.3 Validez externa: Aplicabilidad en contextos reales.

La validez externa se refiere a la generalización de resultados a otras poblaciones o entornos. Factores clave:

- Población:
 - Criterios restrictivos (ej.: exclusión de adultos mayores) limitan aplicabilidad.
- Contexto:
 - ECA en entornos controlados pueden sobrestimar eficacia vs. práctica real.
- Resultados medidos:
 - Mejoras en marcadores sustitutos (ej.: HbA1c) no siempre se traducen en beneficios clínicos tangibles (ej.: prevención de ceguera en diabetes).

Nota: Un análisis detallado de los aspectos anteriores excede el alcance de este documento, por lo que se recomienda consultar documentos específicos para su evaluación integral.

https://revistahigienistas.com/wp-content/uploads/2022/06/Como-leer-un-articulo-cientifico-Las-bases-de-la-medicina-basada-en-la-evidencia-Trisha-Greenhalgh-z-lib.org_.pdf

La lectura crítica de estudios médicos implica dos preguntas clave: ¿El estudio se realizó correctamente? y ¿Sus resultados son útiles para mis pacientes? Al evitar depender solo de

cifras como el valor de p y enfocarse en el impacto real, los médicos pueden tomar decisiones más equilibradas, combinando guías generales con el contexto único de cada caso.

Tras la culminación de la estrategia de búsqueda sistematizada, se efectuó una revisión exhaustiva de los 39 registros documentales identificados. Del corpus total, cinco documentos correspondientes a metaanálisis que abordaban directamente la pregunta de investigación planteada fueron seleccionados para su recuperación en formato de texto completo, con el propósito de proceder a su evaluación crítica detallada.

Listado de documentos recuperados en texto completo.

- i. Drevin G, Andersson B, Svensson JF. Thoracoscopy or Thoracotomy for Esophageal Atresia: A Systematic Review and Meta-analysis. *Ann Surg.* 2021 Dec 1;274(6):945-953. doi: 10.1097/SLA.0000000000004239. PMID: 33002943.
- ii. Way C, Wayne C, Grandpierre V, Harrison BJ, Travis N, Nasr A. Thoracoscopy vs. thoracotomy for the repair of esophageal atresia and tracheoesophageal fistula: a systematic review and meta-analysis. *Pediatr Surg Int.* 2019 Nov;35(11):1167-1184. doi: 10.1007/s00383-019-04527-9. Epub 2019 Jul 29. PMID: 31359222.
- iii. Borruto FA, Impellizzeri P, Montalto AS, Antonuccio P, Santacaterina E, Scalfari G, Arena F, Romeo C. Thoracoscopy versus thoracotomy for esophageal atresia and tracheoesophageal fistula repair: review of the literature and meta-analysis. *Eur J Pediatr Surg.* 2012 Dec;22(6):415-9. doi: 10.1055/s-0032-1329711. Epub 2012 Nov 21. PMID: 23172569.
- iv. Yang YF, Dong R, Zheng C, Jin Z, Chen G, Huang YL, Zheng S. Outcomes of thoracoscopy versus thoracotomy for esophageal atresia with tracheoesophageal fistula repair: A PRISMA-compliant systematic review and meta-analysis. *Medicine (Baltimore).* 2016 Jul;95(30): e4428. doi: 10.1097/MD.0000000000004428. PMID: 27472740; PMCID: PMC5265877.
- v. Wu Y, Kuang H, Lv T, Wu C. Comparison of clinical outcomes between open and thoracoscopic repair for esophageal atresia with tracheoesophageal fistula: a systematic review and meta-analysis. *Pediatr Surg Int.* 2017 Nov;33(11):1147-1157. doi: 10.1007/s00383-017-4153-9. Epub 2017 Sep 15. PMID: 28914345.

Posteriormente, se solicitó a un modelo de IA la aplicación de la herramienta AMSTAR 2 para evaluar la calidad metodológica de cada uno de los cinco estudios seleccionados. A modo de ejemplo, en el presente documento solo se presenta el resultado correspondiente al estudio de Drevin et al. 2021, aunque el mismo procedimiento fue aplicado a todos los artículos.

Investigador: "Adjunto a continuación un archivo adicional. Aplica la herramienta AMSTAR 2 para evaluar la calidad metodológica de la revisión sistemática o metaanálisis incluido y presenta el resultado correspondiente de forma clara y estructurada. Finalmente, basándote en la evaluación de estos ítems, proporciona una calificación global de la calidad metodológica del estudio de acuerdo con los criterios de la herramienta AMSTAR 2."

Respuesta del Modelo de IA:

Evaluación de la Calidad Metodológica mediante AMSTAR 2 de Drevin et al. 2021.

Ítem	Descripción	Cumplimiento	Comentarios
1	Pregunta PICO explícita	✓	Definida claramente: población (RN con atresia), intervención (toracoscopia), comparación (toracotomía), resultados (seguridad/eficacia).
2	Protocolo pre-registrado y justificación de desviaciones	✓	Registrado en PROSPERO (#CRD42019121862). No se mencionan desviaciones.
3	Diseños de estudio justificados	✓	Incluyeron estudios prospectivos/retrospectivos. Excluyeron casos clínicos, revisiones, etc.
4	Estrategia de búsqueda comprehensiva	✓	Búsqueda en PubMed, EMBASE, Web of Science, Cochrane. Términos MeSH/texto libre.
5	Selección y extracción de datos en duplicado	✓	Dos investigadores independientes, con arbitraje de un tercero.
6	Lista de estudios excluidos con justificación	✗	Flujo PRISMA muestra exclusiones, pero sin lista detallada en suplementos.
7	Características de los estudios detalladas	✓	Tablas 1 y 2 resumen diseños, países, tamaños muestrales y datos demográficos.
8	Evaluación del riesgo de sesgo en estudios incluidos	⚠	Usaron herramienta Joanna Briggs, pero sin detallar resultados por estudio.
9	Fuentes de financiamiento de estudios primarios	✗	No se menciona el financiamiento de los estudios primarios.
10	Métodos estadísticos apropiados	✓	Modelos de efectos aleatorios, SMD/OR con IC 95%, conversión de medianas a medias (Hozo et al.).
11	Impacto del riesgo de sesgo en resultados	✗	No discuten cómo el riesgo de sesgo afectó los resultados globales.
12	Consideración del riesgo de sesgo en la interpretación	⚠	Mencionan limitaciones (heterogeneidad), pero sin vincularlo al riesgo de sesgo.
13	Discusión de heterogeneidad clínica/metodológica	✓	Analizan diferencias en definiciones de complicaciones y diseños.
14	Evaluación del sesgo de publicación	⚠	Usaron funnel plots, pero reconocen limitaciones para SMD.
15	Declaración de conflictos de interés	✓	Autores declaran no tener conflictos. No evalúan COI en estudios primarios.
16	Financiamiento de la revisión	✓	Financiado por fundaciones, sin influencia en el estudio.

Calificación global: El metanálisis presenta **calidad metodológica moderada** según AMSTAR 2.

Calidad metodológica del resto de documentos:

- I. Way et al. 2019: Moderada.
- II. Boruruto et al. 2012: Baja.
- III. Yang et al, 2016: Baja.
- IV. Wu et al. 2017: Críticamente baja.

No obstante, la capacidad técnica de solicitar a los modelos de IA la generación de resúmenes estructurados de artículos que contemplen población, intervención, comparador y resultados, se enfatiza al lector la importancia de ejecutar personalmente dicho proceso analítico, considerando las potenciales imprecisiones inherentes a los modelos computacionales actuales.

Si bien las herramientas de IA pueden constituir un apoyo en fases preliminares del proceso de investigación, se recomienda imperativamente que la elaboración de resúmenes de literatura científica sea realizada directamente por el investigador responsable durante la fase de lectura crítica de documental. Este procedimiento metodológico salvaguarda el control integral sobre la exactitud de la información, la adherencia a principios éticos y la contextualización adecuada del contenido científico.

Razones para priorizar la autoría humana:

- Prevención de omisiones críticas:
 - Solo expertos identifican elementos clave (población, intervención, comparadores).
- Manejo de matices metodológicos:
 - La IA simplifica diseños complejos (ej.: "doble ciego" vs. "cruzado").
- Contextualización crítica:
 - Enfoque en impacto real, no en patrones estadísticos automatizados.

Errores comunes con IA.

- Despersonalización del contenido: Pérdida de estilo y tono único.
- Resúmenes genéricos: Sin adaptación al público objetivo (clínico, académico, técnico).
- Omisión de controversias: Limitaciones o debates no reflejados.
- Sesgos algorítmicos: Replicación de inconsistencias en datos de entrenamiento.

Recomendaciones para el uso de IA

- Como apoyo para:
 1. Verificar coherencia entre secciones.
 2. Identificar términos clave (ej.: PICO).
 3. Sugerir estructuras base (para ajustes manuales).

Se presenta a continuación el compendio analítico del estudio desarrollado por Drevin et al. (2021), cuya redacción final fue sometida a un proceso de revisión estilística mediante la implementación de un modelo de IA, con el objetivo de optimizar su estructura y claridad expositiva.

En un metanálisis realizado por de Drevin et al. (que incluyó 17 estudios: 16 observacionales y 1 ECA con un total de 1043 participantes con atresia esofágica [AE]), se comparó la reparación quirúrgica mediante toracotomía (695 participantes) frente a toracoscopia, (348 participantes) para evaluar seguridad y eficacia. Los resultados reportados fueron:

- Menor estancia intrahospitalaria en el grupo de toracoscopia: diferencia de medias estandarizada (DME) = -11.9 días (IC 95%: -23.49 a 0.32).
- Menor tiempo hasta la extubación en el grupo de toracoscopia: DME = -3.22 días (IC 95%: -5.93 a -0.51).
- Menor tiempo hasta la primera alimentación oral en el grupo de toracoscopia: DME = -2.84 días (IC 95%: 0.34 a 4.16).

Sin embargo, la toracoscopia se asoció con:

- Mayor tiempo quirúrgico: DME = +27.69 minutos (IC 95%: 12.6 a 43.32).
- Mayor probabilidad de cirugías antirreflujo adicionales: razón de momios (OR) = 2.2 (IC 95%: 1.06 a 4.42).

No se encontró diferencia estadística entre los grupos en:

- Fugas posquirúrgicas: OR = 1.92 (IC 95%: 0.97 a 3.8).
- Formación de estenosis: OR = 2.66 (IC 95%: 0.86 a 3.23).
- Necesidad de dilatación de estenosis: OR = 1.9 (IC 95%: 0.16 a 3.88).
- Mortalidad: OR = 1.18 (IC 95%: 0.34 a 4.16).

Limitaciones:

- No se reportó la heterogeneidad de los resultados en ningún análisis.
- Calidad metodológica (AMSTAR 2): moderada.

Ejemplo aplicado completo:

Tema: Atresia de esófago: Técnica quirúrgica óptima.

- Pregunta PICO: En recién nacidos con atresia de esófago, ¿el abordaje toracoscópico mínimamente invasivo comparado con la toracotomía es más seguro y eficaz?
- Estudio para responder la pregunta: MET.

Algoritmo de PubMed (fecha de búsqueda: 020425):

("Esophageal Atresia"[MeSH] OR "esophageal atresia"[tiab] OR "oesophageal atresia"[tiab] OR "atresia of esophagus"[tiab] OR "atresia of oesophagus"[tiab] OR "esophageal gap"[tiab] OR "esophageal defect"[tiab] OR "TEF"[tiab] OR "tracheoesophageal fistula"[tiab] OR "tracheo-esophageal fistula"[tiab] OR "TOF"[tiab]) AND ("Thoracoscopy"[MeSH] OR "Thoracic Surgery, Video-Assisted"[MeSH] OR thoracoscop*[tiab] OR "minimally invasive"[tiab] OR "video-assisted"[tiab] OR "endoscopic surgery"[tiab] OR VATS[tiab] OR "minimal access"[tiab] OR "laparoscopic approach"[tiab] OR "laparoscopic repair"[tiab] OR "endoscopic repair"[tiab]) AND ("Thoracotomy"[MeSH] OR thoracotom*[tiab] OR "open surgery"[tiab] OR "open approach"[tiab] OR "open repair"[tiab] OR "conventional surgery"[tiab] OR "traditional approach"[tiab] OR "surgical repair"[tiab] OR "standard repair"[tiab]) AND ("Meta-Analysis"[pt] OR "Meta-Analysis as Topic"[MeSH] OR metaanaly*[tiab] OR meta-analy*[tiab] OR "meta analysis"[tiab] OR "systematic review"[tiab] OR "systematic overview"[tiab] OR "systematic literature review"[tiab] OR "Systematic Review"[Publication Type] OR "Review Literature as Topic"[MeSH] OR "Review"[Publication Type] OR "Cochrane Database Syst Rev"[Journal] OR "umbrella review"[tiab] OR "meta-synthesis"[tiab] OR metasynthesis[tiab] OR "integrative review"[tiab])

- Resultados totales: 39.
- Resultados potencialmente útiles: 5.
- Lectura crítica de un documento obtenido de la búsqueda:
 - En un metanálisis realizado por de Drevin et al. (que incluyó 17 estudios: 16 observacionales y 1 ECA con un total de 1043 participantes con atresia esofágica [AE]), se comparó la reparación quirúrgica mediante toracotomía (695 participantes) frente a toracoscopia, (348 participantes) para evaluar seguridad y eficacia. Los resultados reportados fueron:
 - Menor estancia intrahospitalaria en el grupo de toracoscopia: diferencia de medias estandarizada (DME) = -11.9 días (IC 95%: -23.49 a 0.32).
 - Menor tiempo hasta la extubación en el grupo de toracoscopia: DME = -3.22 días (IC 95%: -5.93 a -0.51).
 - Menor tiempo hasta la primera alimentación oral en el grupo de toracoscopia: DME = -2.84 días (IC 95%: 0.34 a 4.16).
 - Sin embargo, la toracoscopia se asoció con:
 - Mayor tiempo quirúrgico: DME = +27.69 minutos (IC 95%: 12.6 a 43.32).
 - Mayor probabilidad de cirugías antirreflujo adicionales: razón de momios (OR) = 2.2 (IC 95%: 1.06 a 4.42).
 - No se encontró diferencia estadística entre los grupos en:
 - Fugas posquirúrgicas: OR = 1.92 (IC 95%: 0.97 a 3.8).
 - Formación de estenosis: OR = 2.66 (IC 95%: 0.86 a 3.23).
 - Necesidad de dilatación de estenosis: OR = 1.9 (IC 95%: 0.16 a 3.88).
 - Mortalidad: OR = 1.18 (IC 95%: 0.34 a 4.16).
 - Limitaciones:
 - No se reportó la heterogeneidad de los resultados en ningún análisis.
 - Calidad metodológica (AMSTAR 2): moderada.

Nota: La aplicación de evidencia debe considerar experiencia clínica, recursos disponibles y preferencias del paciente. Este manual no sustituye el juicio profesional.

6. INTELIGENCIA ARTIFICIAL EN MEDICINA BASADA EN EVIDENCIA: LIMITACIONES Y DESAFÍOS ÉTICOS.

La integración de IA en la MBE ofrece ventajas significativas, pero su aplicación conlleva limitaciones y dilemas éticos críticos que todo profesional debe considerar:

6.1 Limitaciones técnicas de la IA.

- Falta de comprensión contextual:
 - Los modelos de IA pueden confundir términos similares (ej.: "estenosis" vs. "esclerosis") o ignorar matices en diseños de estudio (ej.: diferencias entre cohortes prospectivas y retrospectivas).
 - Consecuencia: Algoritmos con alta sensibilidad, pero baja especificidad.
- Sesgo en datos de entrenamiento:
 - Modelos entrenados con literatura en inglés o de países de altos ingresos perpetúan inequidades al excluir evidencia regional (ej.: bases como LILACS).
- Sobreajuste y sobregeneralización:
 - Sobreajuste: Ecuaciones booleanas demasiado específicas que omiten estudios con terminología alternativa.
 - Sobregeneralización: Estrategias amplias que recuperan literatura irrelevante.
- Falta de transparencia en modelos propietarios:
 - Plataformas comerciales operan como "cajas negras", dificultando la evaluación crítica de su rigor metodológico.

6.2 Desafíos éticos.

- Sesgo algorítmico y justicia social:
 - Algoritmos que priorizan terapias costosas (comunes en naciones con altos niveles de desarrollo y capacidad de inversión en salud) agravan desigualdades en entornos de bajos recursos.
- Privacidad y consentimiento:
 - Uso de datos de pacientes o estudios sin consentimiento explícito (ej.: "raspado" de historias clínicas).
- Responsabilidad legal:
 - Vacíos en marcos legales para atribuir culpas por decisiones erróneas basadas en IA.
- Erosión de la autonomía clínica:
 - Dependencia excesiva en recomendaciones automatizadas, reduciendo el pensamiento crítico.
- Falta de transparencia:

- Modelos comerciales ocultan metodologías por "*secretos industriales*", contradiciendo el principio de reproducibilidad científica.

6.3 Recomendaciones para un uso ético:

- Verifique fuentes primarias: No confíe en resúmenes automatizados sin revisar los estudios originales.
- Transparencia: Acceso a ecuaciones booleanas y criterios de selección de términos.
- Combine IA con *expertise* humano: Consulte bibliotecólogos o metodólogos clínicos.

La integración de inteligencias artificiales en la medicina basada en evidencia no solo implica la optimización de procesos, sino también la adopción de un marco ético riguroso y una total transparencia. Más allá de las limitaciones técnicas de los modelos de caja negra, la necesidad de inteligencia artificial explicable (Explainable AI - XAI) en la investigación médica es crucial. En un ámbito donde las decisiones impactaran directamente la salud de los pacientes o la toma de decisiones administrativas, no basta con que un algoritmo genere un resultado; es imperativo comprender cómo llegó a esa conclusión. Esto permite a los profesionales auditar las decisiones, identificar posibles sesgos inherentes al entrenamiento del modelo y garantizar que la lógica subyacente sea médicamente sólida y justa.

Para asegurar una IA responsable, es fundamental promover:

- Rendición de Cuentas: Establecer quién es responsable por los resultados derivados del uso de la IA.
- Equidad y justicia social: Asegurar que los algoritmos no perpetúen ni agraven desigualdades existentes, por ejemplo, al priorizar terapias costosas o excluir evidencia regional.
- Seguridad y fiabilidad: Garantizar que los modelos de IA sean robustos y sus resultados consistentes.
- Privacidad y Consentimiento: Respetar las normativas de protección de datos cuando se utilicen datos de pacientes para el entrenamiento o validación de modelos.

Además, la transparencia en la publicación de datos y decisiones es un pilar esencial. A diferencia de algunas herramientas de IA que operan como cajas negras, algunas herramientas como "Gemini" permiten la extracción de un archivo de datos que proporciona una visión detallada de todas las decisiones tomadas durante el proceso asistido por IA. Se recomienda que los investigadores publiquen estos archivos de datos junto con sus revisiones sistemáticas para garantizar la reproducibilidad, aumentar la fiabilidad y fomentar la confianza en los resultados asistidos por IA.

Esta práctica de apertura metodológica es clave para que la inteligencia artificial se integren plenamente en la práctica de MBE con el máximo rigor científico y ético.

7. RECURSOS TECNOLÓGICOS EMERGENTES EN MBE.

7.1 Large Language Models (LLMs) Especializados.

- Ejemplos:
 - EBM-LLM: Genera resúmenes estructurados (PICO) a partir de abstracts, entrenado en revisiones Cochrane y guías NICE.
 - Cochrane AI: Reduce un 40% el tiempo de selección de estudios para actualizar revisiones sistemáticas.
- Ventajas:
 - Aceleran la identificación de evidencia en crisis sanitarias (ej.: pandemias).
 - Reducen carga laboral en metaanálisis en red.
- Desafíos:
 - Alucinaciones: Referencias inexistentes o malinterpretación de resultados.
 - Acceso limitado: Desarrollados por empresas privadas, restringiendo su uso en entornos sin recursos.

7.2 Repositorios de Real-World Data (RWD).

- Ejemplos de plataformas:
 - TriNetX: Analiza efectividad de tratamientos en 250 millones de pacientes mediante historias clínicas electrónicas (HCEs).
 - FDA Sentinel: Monitoriza seguridad de medicamentos con datos de seguros y hospitales.
 - OMOP Common Data Model: Estandariza datos heterogéneos para comparaciones internacionales.
- Aplicaciones:
 - Detección de efectos adversos raros (ej.: miocarditis post-vacunas COVID-19).
 - Evaluación de costo-efectividad en poblaciones específicas (ej.: ancianos con comorbilidades).
- Limitaciones:
 - Sesgo de selección: Exclusión de poblaciones vulnerables.
 - Calidad variable: Errores en codificación diagnóstica o falta de estandarización en HCEs.

La innovación tecnológica en MBE es indispensable, pero requiere:

- Validación rigurosa: Evitar conclusiones erróneas por limitaciones de IA o RWD.
- Transparencia y ética: Garantizar que los avances técnicos sirvan a la equidad en salud.

8. CONCLUSIONES.

La medicina basada en evidencia (MBE) y la inteligencia artificial (IA) se han consolidado como pilares fundamentales para transformar la práctica clínica, la investigación y la toma de decisiones institucionales. Este manual destaca que la búsqueda sistematizada no es solo un ejercicio técnico, sino un proceso riguroso que te exige precisión metodológica, adaptabilidad y un dominio profundo de las herramientas disponibles. La integración de IA en la creación de algoritmos de búsqueda, el uso de tesauros (MeSH, Emtree, DeCS) y la aplicación de operadores booleanos no solo optimizan la recuperación de evidencia relevante, sino que también mitigan sesgos como la omisión de literatura gris o las limitaciones idiomáticas.

La pregunta PICO se posiciona como el eje central para guiar tu búsqueda, asegurando que cada estrategia se alinee con las necesidades clínicas específicas. Sin embargo, su máximo potencial se alcanza al identificar correctamente el tipo de estudio adecuado (ECA para intervenciones, cohortes para pronósticos) y respetar la jerarquía de la evidencia científica. Este enfoque garantiza que tus conclusiones se sustenten en diseños robustos, minimizando el riesgo de basar decisiones en hallazgos espurios o de baja calidad metodológica.

La lectura crítica, por su parte, te exige ir más allá del valor de *p* para evaluar el impacto clínico real. Herramientas como RoB 2.0 te permiten distinguir entre estudios válidos y aquellos con limitaciones, mientras que la validez externa asegura que los hallazgos puedan aplicarse al contexto particular de tu paciente o institución. En este sentido, la IA no reemplaza tu juicio clínico, sino que potencia tu capacidad para sintetizar información compleja, priorizar evidencia y reducir el tiempo en revisiones manuales.

Desafíos persistentes:

Las revisiones sistemáticas, aunque son herramientas indispensables para sintetizar la evidencia, corren el riesgo de quedar desactualizadas rápidamente, a menudo en pocos años, debido a la continua publicación de nuevas investigaciones. La capacidad de la IA para acelerar la revisión de literatura permite realizar actualizaciones más frecuentes y menos laboriosas de las revisiones sistemáticas. Esto asegura que las conclusiones y recomendaciones se basen en la evidencia más reciente disponible, contribuyendo a una práctica clínica y políticas de salud más informadas y relevantes.

El equilibrio entre sensibilidad y especificidad en las búsquedas sigue siendo un acto de precisión: estrategias demasiado amplias generan "ruido", mientras que las restrictivas pueden omitir estudios clave o provocar silencio documental.

La rápida evolución de la IA te obliga a mantenerte actualizado, tanto en el manejo de nuevas plataformas (metabuscadores, repositorios de *big data*) como en la evaluación crítica de algoritmos automatizados.

En un panorama donde la información biomédica crece exponencialmente, este manual refuerza que la excelencia en la MBE reside en la sinergia entre tecnología y *expertise* humano. La IA acerca la evidencia al punto de atención, pero es tu formación metodológica, curiosidad



científica y capacidad para contextualizar hallazgos lo que transforma datos en acciones clínicas significativas. Así, se consolida un ciclo virtuoso: preguntas mejor estructuradas, búsquedas más eficientes y decisiones informadas que elevan la calidad asistencial, la investigación y, en última instancia, los resultados en salud.

La IA y la MBE coexisten en un ecosistema en constante cambio. La sinergia entre una tecnología en evolución rápida y la experiencia humana, con su juicio crítico y capacidad de contextualización, es fundamental para transformar eficazmente la información en acciones clínicas significativas y mantener la relevancia del conocimiento en un entorno médico en constante expansión.

9. ANEXOS.

9.1 Uso de la Interfaz de Chat con IA.

La interfaz de chat con IA es una herramienta digital que permite a los usuarios interactuar en tiempo real con un modelo de IA. A través de esta plataforma, es posible realizar preguntas, obtener explicaciones, generar contenido, analizar documentos y mucho más. Está diseñada para ofrecer una experiencia fluida e intuitiva, similar a una conversación en aplicaciones de mensajería instantánea.

Componentes principales de la interfaz.

Ver imagen 3: "Interfaz de Chat con IA"

- Ventana de chat: Muestra el intercambio de mensajes entre el usuario y la IA.
- Campo de texto: Área donde el usuario escribe sus preguntas, instrucciones o solicitudes.
- Botón de envío: Envía el mensaje escrito para que la IA lo procese.
- Botón de cargar archivo: Permite adjuntar documentos (PDF, Word, imágenes, hojas de cálculo, etc.) para su análisis.

Cómo redactar instrucciones efectivas.

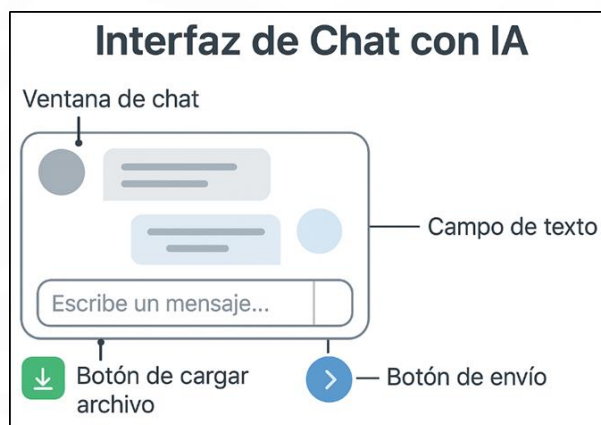
Para obtener respuestas útiles, claras y relevantes de la IA, se recomienda lo siguiente:

Utiliza un lenguaje claro y específico. Ejemplo: "Resume los hallazgos principales de este artículo científico."

Indica el propósito de la consulta si es relevante. Ejemplo: "Necesito esta información para una revisión sistemática."

Si adjuntas un archivo, menciona su contenido y lo que deseas obtener. Ejemplo: "He adjuntado un artículo. ¿Podrías extraer los criterios de inclusión y exclusión que menciona?"

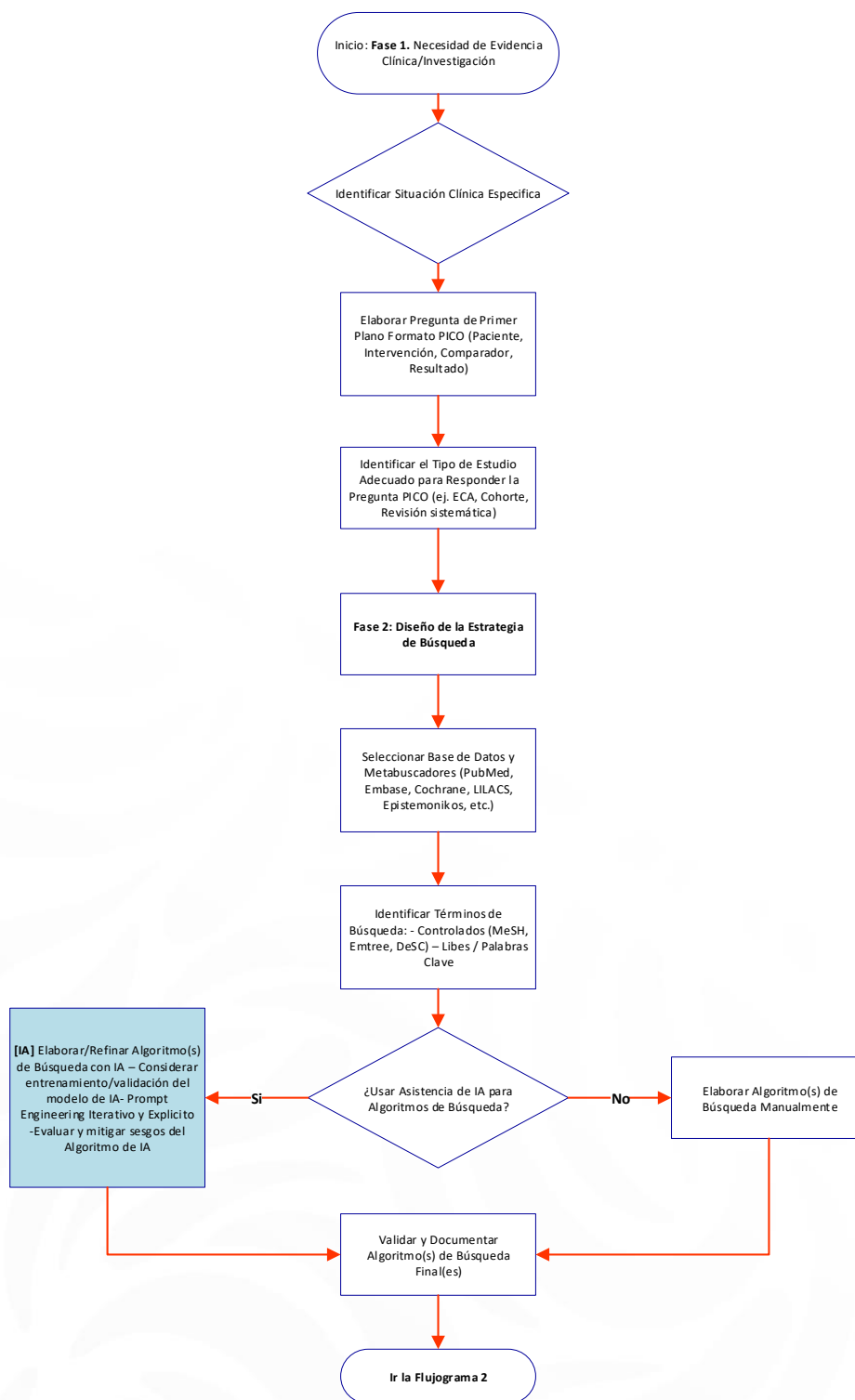
Imagen 3.



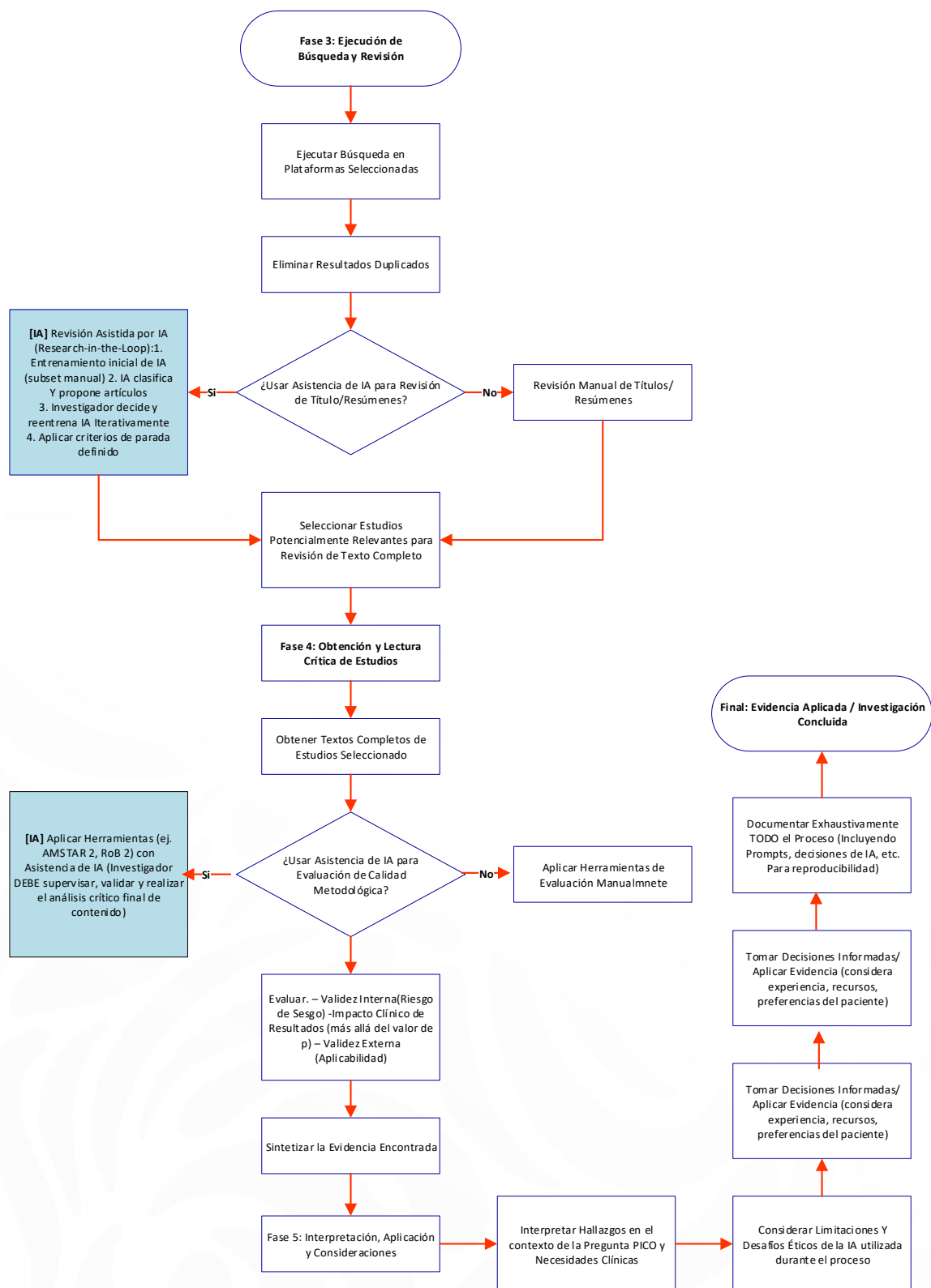
Fuente de la imagen: Generada por Claude 4 – Anthropic.

9.2 Flujogramas

Flujograma 1.



Flujograma 2.



9.3 Glosario.

Alucinación en IA: Fenómeno donde modelos de IA generan información falsa o inventada (ej.: referencias bibliográficas inexistentes) debido a limitaciones en su entrenamiento o sesgos en los datos.

AMSTAR 2: Herramienta validada para evaluar la calidad metodológica de revisiones sistemáticas y metaanálisis. Incluye 16 ítems que valoran protocolo, selección de estudios, evaluación de sesgos y análisis estadísticos.

Búsqueda narrativa: Enfoque informal de búsqueda bibliográfica basado en términos intuitivos, sin estructura rigurosa ni documentación detallada. Útil para exploraciones iniciales.

Búsqueda sistematizada: Proceso estructurado y reproducible que combina términos controlados, vocabulario libre y operadores booleanos para minimizar sesgos y garantizar exhaustividad en la recuperación de evidencia.

Cochrane AI: Herramienta desarrollada por la Colaboración Cochrane que emplea IA para identificar estudios relevantes y actualizar revisiones sistemáticas de manera semi-automatizada.

Data Scientist: Profesional especializado en analizar grandes volúmenes de datos, clave para interpretar real-world data o entrenar modelos de IA en MBE.

EBM-LLM: Large Language Model (LLM) especializado en Medicina Basada en Evidencia, entrenado para sintetizar hallazgos científicos y generar resúmenes estructurados.

Ecuaciones booleanas: Expresiones lógicas que combinan términos de búsqueda mediante operadores booleanos (AND, OR, NOT) para recuperar literatura científica de manera estructurada y precisa.

Ensayo Clínico Aleatorizado: Diseño de investigación que asigna participantes aleatoriamente a grupos de intervención o control, considerado el estándar de oro para evaluar tratamientos.

Estudios de casos y controles: Diseño observacional que compara sujetos con una condición (casos) frente a otros sin ella (controles), para identificar asociaciones con exposiciones previas.

Estudios de cohorte: Estudio observacional que sigue grupos expuestos y no expuestos a un factor para analizar la incidencia de desenlaces. Puede ser prospectivo o retrospectivo.

Estudios observacionales: Investigaciones que analizan relaciones sin intervención activa, incluyendo cohortes, casos y controles, o series de casos.

FDA Sentinel: Sistema de vigilancia de la FDA que utiliza real-world data (RWD) para monitorizar la seguridad de medicamentos y dispositivos médicos.

GDPR (Reglamento General de Protección de Datos): Normativa europea que regula el manejo de datos personales, aplicable al uso de IA con información de pacientes.

Heterogeneidad: Variabilidad en resultados entre estudios incluidos en un metaanálisis, originada por diferencias metodológicas, poblacionales o en intervenciones.

HIPAA (Ley de Portabilidad y Responsabilidad de Seguros Médicos): Legislación estadounidense que protege la privacidad de datos de salud, relevante al usar IA con historias clínicas.

Historias Clínicas Electrónicas (HCEs): Registros digitales de pacientes que almacenan diagnósticos, tratamientos y resultados, fuente clave de real-world data.

Literatura gris: Documentos científicos no publicados en revistas arbitradas, como tesis, informes técnicos o registros de ensayos clínicos.

LLM (Large Language Model): Modelo de IA entrenado con vastos corpus de texto para generar, resumir o analizar lenguaje natural (ej.: GPT-4, LLaMA).

Medicina Basada en Evidencia: Enfoque clínico que integra la mejor evidencia disponible, la experiencia profesional y las preferencias del paciente.

Metabuscador: Plataforma que integra resultados de múltiples bases de datos simultáneamente (ej.: Epistemonikos, Tripdatabase), optimizando la cobertura de la evidencia.

Metanálisis en red: Técnica estadística que compara múltiples intervenciones indirectamente mediante análisis combinados de estudios existentes.

Modelo de inteligencia artificial: Algoritmo computacional entrenado para reconocer patrones en datos, realizar predicciones o generar respuestas automatizadas. En el contexto de este manual, se refiere a sistemas que apoyan la creación de algoritmos de búsqueda, validación de estrategias o análisis crítico de estudios (ej.: generación de ecuaciones booleanas para PubMed o Embase).

Motor de búsqueda: Herramienta especializada que indexa literatura en una base específica (ej.: PubMed para MEDLINE), con filtros avanzados por tipo de estudio o idioma.

OMOP Common Data Model: Estándar internacional para estructurar datos médicos heterogéneos, facilitando su análisis en investigación traslacional.

OpenGrey (antes conocido como SIGLE - System for Information on Grey Literature in Europe) es una base de datos especializada en literatura gris de acceso abierto.

PMID (PubMed Identifier): Número único asignado a cada artículo en PubMed, usado para verificar referencias generadas por IA.

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA): Conjunto de estándares para reportar revisiones sistemáticas y metaanálisis, asegurando transparencia.

Prompts: Instrucciones o consultas ingresadas a un modelo de IA para guiar su respuesta (ej.: "Genera una ecuación booleana para PubMed sobre diabetes tipo 2").

QUADAS-2: Herramienta para evaluar riesgo de sesgo en estudios de pruebas diagnósticas, considerando selección de participantes, estándar de referencia y flujo de pruebas.



Real-World Data (RWD): Datos obtenidos de fuentes no controladas (historias clínicas, registros de seguros), útiles para estudiar efectividad en práctica clínica real.

Revisiones sistemáticas: Síntesis rigurosa de evidencia que aplica métodos explícitos para identificar, seleccionar y analizar estudios relevantes.

Risk of Bias 2.0 (RoB2): Instrumento de Cochrane para evaluar riesgo de sesgo en ECA, analizando dominios como aleatorización y enmascaramiento.

Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I): Herramienta para valorar sesgos en estudios no aleatorizados, como cohortes o casos y controles.

STrengthening the Reporting of OBservational studies in Epidemiology (STROBE): Lista de verificación para mejorar la transparencia en la publicación de estudios observacionales.

Sesgo algorítmico: Distorsión sistemática en los resultados de IA debido a datos de entrenamiento no representativos o criterios de diseño discriminatorios.

Sesgo de base de datos: Limitación al buscar evidencia solo en plataformas específicas, omitiendo estudios relevantes indexados en otras fuentes.

Sesgo de publicación: Distorsión en la evidencia disponible debido a la mayor probabilidad de publicar estudios con resultados positivos o significativos.

Sesgo de selección: Error sistemático que surge cuando los participantes del estudio no representan adecuadamente a la población objetivo.

Series de casos: Diseño observacional que describe características o resultados de un grupo de pacientes con una condición común, sin grupo de comparación.

Tesoros en ciencias de la salud: Vocabularios controlados que unifican y agrupa términos médicos (ej.: MeSH en PubMed) para estandarizar búsquedas bibliográficas.

TriNetX: Plataforma global que agrega real-world data de millones de pacientes para investigación clínica y análisis de efectividad terapéutica.

Validez externa: Grado en que los resultados de un estudio pueden generalizarse a otras poblaciones, contextos o entornos clínicos.

Validez interna: Rigor metodológico de un estudio para establecer relaciones causales confiables, minimizando errores sistemáticos.

WHO ICTRP (Plataforma de Registros Internacionales de Ensayos Clínicos de la OMS): Repositorio que centraliza ensayos clínicos registrados en todo el mundo, clave para evitar sesgos de publicación.

10. BIBLIOGRAFÍA.

- Alarcón Palacios Marco, Ojeda Gómez Roberto Carlos, Ticse Huaricanha Iveht Lucy, & Cajachagua Hilario Catherine. (2015). Análisis crítico de ensayos clínicos aleatorizados: Riesgo de sesgo. *Rev Estomatol Herediana.*, 25(4), 304–308.
- Aletti Federico. (2015, July). A Comparison of Emtree® and MeSH. Elsevier R&D Solutions.
- Anza Aguirrezavala ignacio, Lopez Santiago Asencio, Gonzalez Valverde Francisco Miguel, Ruiz Merino Guadalupe, Garcia Teresa, Valverde Iniesta Juan Jose, & Cinesi Gomez Cesar. (n.d.). METODOLOGÍA DE LA INVESTIGACIÓN Y PRÁCTICA CLÍNICA BASADA EN LA EVIDENCIA. PROGRAMA TRANSVERSAL Y COMPLEMENTARIO DEL RESIDENTE (PTCR). In consejería de murcia. Quaderna Editorial.
- Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. *J Med Libr Assoc.* 2018 Oct 1;106(4):531-541.
- Cediél-Becerra Ntalia M, & Krause Gerard. (2013). Herramientas para la toma de decisiones en salud pública basadas en la evidencia y priorización de enfermedades. *Rev. Salud Pública.*, 15(5), 694–706.
- Centro de innovación, ministerio de educación. (2025, May 6). Guía para Docentes: Cómo usar ChatGPT para potenciar el aprendizaje activo. Ministerio de Educación de Chile.
- Duarte Martínez, M. C. (2019). Medicina basada en la evidencia: búsqueda en bases de datos en internet y términos mesh. *Medicina*, 19(4), 218–222. <https://doi.org/10.23878/medicina.v19i4.732>.
- Gisbert, J. P., & Bonfill, X. (2004). ¿Cómo realizar, evaluar y utilizar revisiones sistemáticas y metaanálisis? *Gastroenterología y Hepatología*, 27(3), 129–149. [https://doi.org/10.1016/S0210-5705\(03\)79110-9](https://doi.org/10.1016/S0210-5705(03)79110-9).
- González-Padilla, D. A., & Dahm, P. (2023). Evaluating the Certainty of Evidence in Evidence-based Medicine. *European Urology Focus*, 9(5), 708–710. <https://doi.org/10.1016/j.euf.2023.10.014>.
- INSTITUTO DE CIENCIAS DE LA SALUD. (2012, May 6). Guía para hacer búsquedas bibliográficas. INSTITUTO DE CIENCIAS DE LA SALUD Biblioteca.
- Jayakumar, S., Sounderajah, V., Normahani, P., Harling, L., Markar, S. R., Ashrafian, H., & Darzi, A. (2022). Quality assessment standards in artificial intelligence diagnostic accuracy systematic reviews: a meta-research study. *Npj Digital Medicine*, 5(1), 11. <https://doi.org/10.1038/s41746-021-00544-y>.
- Khosravi, M., Zare, Z., Mojtabaieian, S. M., & Izadi, R. (2024). Artificial Intelligence and Decision-Making in Healthcare: A Thematic Analysis of a Systematic Review of Reviews. *Health Services Research and Managerial Epidemiology*, 11. <https://doi.org/10.1177/23333928241234863>.
- Lazcano-Ponce, E., Salazar-Martínez, E., Gutiérrez-Castrellón, P., Angeles-Llerenas, A., Hernández-Garduño, A., & Viramontes, J. L. (2004). Ensayos clínicos aleatorizados: variantes, métodos de aleatorización, análisis, consideraciones éticas y regulación.

- Salud Pública de México, 46(6), 559–584. <https://doi.org/10.1590/S0036-36342004000600012>.
- Lazcano-Ponce, E., Salazar-Martínez, E., & Hernández-Avila, M. (2001). [Case-control epidemiological studies: theoretical bases, variants and applications]. *Salud Publica de Mexico*, 43(2), 135–150.
 - Malaga German, & Neira-Sánchez Elsa R. (2018). La medicina basada en la evidencia, su evolución a 25 años desde su diseminación, promoviendo una práctica clínica científica, cuidadosa, afectuosa y humana. *Acta Médica Peruana*, 35(2), 121–126.
 - Martínez Ferreras, D. (s.f.). Los tesauros. Recuperado de: <http://cvapp.uoc.edu/autors/MostraPDFMaterialAction.do?id=143963>.
 - Martínez Juárez Alejandro. (2009). Estudios de cohorte. *Perinatol Reprod Hum*, 23(1), 18–24. <https://www.inper.mx/descargas/pdf/EstudiosdeCohorte.pdf>.
 - Masic, I. (2022). Medical Decision Making. *Acta Informatica Medica*, 30(3), 230. <https://doi.org/10.5455/aim.2022.30.230-235>.
 - Molina Arias M. (2015). El metaanálisis en pruebas diagnósticas. *Rev Pediatr Aten Primaria*, 17, 281–285.
 - Nabzo, S., & Fau, C. (2020). Metaanálisis: bases conceptuales, análisis e interpretación estadística. *Revista Mexicana de Oftalmología*, 94(6S), 260–273. <https://doi.org/10.24875/RMO.M20000134>.
 - Philco Lima Patricia, & Navia Bueno María del Pilar. (2009). Medicina basada en evidencia: Búsqueda sistemática de la literatura médica. *Cuadernos Hospital de Clínicas*, 54(1), 75–77.
 - Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, et al. Oxford Centre for Evidence-Based Medicine - Levels of Evidence (March 2009). Oxford: CEBM; 2009.
 - Rivas-Ruiz, R., Castelán-Martínez, O. D., Pérez, M., & Talavera, J. O. (2013). [Clinical research XVII. χ^2 test, from the expected to the observed]. *Revista Médica Del Instituto Mexicano Del Seguro Social*, 51(5), 552–557.
 - RivasRuiz Rodolfo, Moreno-Palacios Jorge, & Talavare Juan O. (2013). Investigación clínica XVI Diferencias de medianas con la U de Mann-Whitney. *Rev Med Inst Mex Seguro Soc*, 51(4), 414–419.
 - Rivas-Ruiz Rodolfo, Perez-Rodriguez Marcela, & Talavera Juan O. (2013). Investigación clínica XV Del juicio clínico al modelo estadístico. Diferencia de medias. Prueba t de Student. *Rev Med Inst Mex Seguro Soc*, 51(3), 300–303.
 - Senpau Marta Millaret. (2019). Búsqueda sistemática de evidencia (FUOC, Ed.; 1st ed.). FUOC.
 - Soto, M., & Rada, G. (2003). Formulación de Preguntas en Medicina Basada en la Evidencia. *Revista Médica de Chile*, 131(10). <https://doi.org/10.4067/S0034-98872003001000016>.
 - Urrutia, G., Torta, S., & Bonfill, X. (2005). Metaanálisis (QUOROM). *Medicina Clínica*, 125, 32–37. [https://doi.org/10.1016/S0025-7753\(05\)72207-7](https://doi.org/10.1016/S0025-7753(05)72207-7).
 - Zamora J, & Abaira V. (2008). Análisis de la calidad de los estudios de evaluación de pruebas diagnósticas. *Nefrología*, 28(sup 2), 42–45. <file:///Users/joseluisaguiar/Downloads/An%C3%A1lisis%20de%20la%20calidad%20de%20los%20estudios%20de%20evaluaci%C3%B3n%20de%20>