

Inteligencia artificial en la toma de decisiones clínicas: revisión sistemática con síntesis narrativa de la literatura publicada entre 2019 y 2026

Artificial intelligence in clinical decision making: a systematic review with narrative synthesis of the literature published between 2019 and 2026

Noemi Monserrath Matza-Aldaz ^{1*}
Washington Jordano Zambrano-Baque ²

*Correspondencia: nmmatza1996@gmail.com

Recibido: 14/03/2026 | Aceptado: 27/03/2026 | Publicado: 05/04/2026

RESUMEN

Introducción: la inteligencia artificial (IA) se incorpora con rapidez al apoyo de las decisiones clínicas, pero la evidencia sobre su utilidad y sus riesgos está dispersa. **Objetivo:** analizar la influencia de la IA en la toma de decisiones clínicas entre 2019 a marzo de 2026: su utilidad, sus desafíos y oportunidades y su aporte a la atención del paciente. **Metodología:** revisión sistemática con síntesis narrativa en dos fases, según PRISMA 2020. La primera fase (2019 a 2023) buscó en Web of Science, Scopus, PubMed y Springer Link, y la actualización (2024 a marzo de 2026) en el Directorio de Revistas de Acceso Abierto. La calidad se valoró con diez criterios (corte en 7,5 de 10). **Resultados:** se incluyeron 63 registros (62 estudios), 26 de la primera fase y 37 de la actualización. Frente a una primera fase descriptiva, la actualización aportó estudios primarios, entre ellos cinco ensayos aleatorizados y ocho con modelos de lenguaje. La IA mejoró la detección de cáncer, la auscultación y la rapidez de algunas decisiones, pero no siempre el razonamiento diagnóstico, y se documentaron dependencia inadecuada, falsos positivos y pérdida de habilidades. Los profesionales piden validación, transparencia, integración en la historia clínica y claridad sobre la responsabilidad. No hubo estudios latinoamericanos. **Conclusiones:** la IA es un apoyo útil cuando se valida en su contexto y se usa bajo supervisión médica, y su beneficio depende de cómo se integra en el trabajo clínico. Hacen falta ensayos con desenlaces en pacientes y estudios en Latinoamérica.

Palabras clave: inteligencia artificial; toma de decisiones clínicas; sistemas de apoyo a la decisión clínica; modelos de lenguaje de gran tamaño; revisión sistemática; seguridad del paciente.

ABSTRACT

Introduction: artificial intelligence (AI) is rapidly entering clinical decision support, but the evidence on its usefulness and risks is scattered and changes quickly. **Objective:** to analyze the influence of AI on clinical decision making between 2019 to March 2026, focusing on its usefulness, its challenges and opportunities, and its contribution to patient care. **Methods:** two-phase systematic review with narrative synthesis, following PRISMA 2020. The first phase (2019 to 2023) searched Web of Science, Scopus, PubMed and Springer Link, and the update (January 2024 to March 2026) searched the Directory of Open Access Journals, with a complementary search. Quality was appraised with ten criteria (cutoff of 7.5 out of 10). **Results:** 63 records (62 studies) were included, 26 from the first phase and 37 from the update. The first phase gathered mostly descriptive papers. The update contributed primary studies, including five randomized trials and eight studies of large language models. AI improved cancer detection, auscultation and the speed of some decisions, but not always diagnostic reasoning, and inappropriate reliance, false positives and deskilling were documented. Clinicians make their use conditional on validation, transparency, integration into the health record and clarity about accountability. There were no Latin American studies. **Conclusions:** AI is a useful aid to clinical decisions when it is validated in its setting and used under medical supervision, and its benefit depends on how it is integrated into clinical work. Trials with patient outcomes and studies in Latin America are needed.

Keywords: artificial intelligence; clinical decision making; clinical decision support systems; large language models; systematic review; patient safety.

1. INTRODUCCIÓN

La inteligencia artificial (IA) reúne técnicas, como el aprendizaje automático y el aprendizaje profundo, que aprenden patrones a partir de datos y los usan para predecir o clasificar. En medicina se aplica para apoyar decisiones de diagnóstico, pronóstico, tratamiento, seguimiento y prevención. El interés tiene

¹ Médico General, Hospital General IESS, Departamento de Cirugía General, Latacunga-Ecuador, E-mail: nmmatza1996@gmail.com, ORCID: <https://orcid.org/0009-0002-0873-5936>

² Médico General, Universidad Nacional de Chimborazo, Riobamba-Ecuador E-mail: washington.zambrano@unach.edu.ec, ORCID: <https://orcid.org/0000-0003-2174-0959>

una razón práctica, porque el error diagnóstico sigue causando daño evitable. Solo en Estados Unidos, cerca de 800 000 personas sufren cada año muerte o discapacidad permanente por esta causa (Newman-Toker et al., 2024).

La literatura sobre el tema ha crecido con rapidez, pero la evidencia de uso real avanza más despacio. Una revisión de alcance identificó 86 ensayos clínicos aleatorizados sobre IA en la práctica clínica publicados hasta noviembre de 2023. El 81 % informó un resultado principal favorable, aunque el 63 % se realizó en un solo centro y la mayoría midió el rendimiento diagnóstico y no resultados relevantes para el paciente (Han et al., 2024). Otra revisión encontró beneficios para los pacientes en 12 de 19 estudios, pero solo cuatro tenían bajo riesgo de sesgo (Wilhelm et al., 2025). En atención primaria, apenas seis de 421 artículos evaluaron estas herramientas en consultas reales (Gomez-Cabello et al., 2024), y una revisión que incluyó bases latinoamericanas retuvo 10 de 77 artículos (Correia y Valério Netto, 2025).

A partir de 2023, el desarrollo de modelos de lenguaje de gran tamaño (Large Language Models, LLM), como ChatGPT, amplió significativamente las aplicaciones potenciales de la inteligencia artificial en el ámbito sanitario. Aunque estos sistemas han demostrado capacidad para analizar información clínica y generar hipótesis diagnósticas, su rendimiento continúa siendo objeto de evaluación. Un metaanálisis de 83 estudios estimó una exactitud diagnóstica promedio del 52,1 %, sin diferencias significativas respecto a médicos no especialistas y con resultados inferiores a los obtenidos por especialistas. Estos hallazgos respaldan su utilización como herramienta de apoyo a la toma de decisiones clínicas, pero evidencian que aún no alcanzan un nivel de desempeño que permita reemplazar la valoración experta (Takita et al., 2025).

La Organización Mundial de la Salud (OMS) fijó en 2021 seis principios éticos para la IA en salud, entre ellos la autonomía, la transparencia, la rendición de cuentas y la equidad (OMS, 2021), y en 2024 añadió orientaciones para los modelos multimodales de gran tamaño (OMS, 2024). La Unión Europea exige gestión de riesgos, calidad de los datos, transparencia y supervisión humana a los sistemas de alto riesgo (Unión Europea, 2024). La Organización Panamericana de la Salud (OPS) incluyó la IA entre los principios de la transformación digital y llamó a prevenir los sesgos algorítmicos (OPS, 2021), y en Ecuador la ley clasifica los datos de salud como datos sensibles (Asamblea Nacional del Ecuador, 2021). Para la investigación existen guías de reporte específicas, como DECIDE-AI (Vasey et al., 2022) y TRIPOD+AI (Collins et al., 2024).

Una versión previa de esta revisión, limitada a 2019 y 2023, encontró una literatura sobre todo descriptiva, con muy pocos datos medidos. Dada la velocidad del campo, era necesario actualizarla hasta 2026 para saber si esa situación cambió. La pregunta principal fue cómo ha influido la IA en la toma de decisiones clínicas en medicina (PP1). Tres preguntas específicas la desarrollaron: (PS1) cuál es la utilidad de la IA para mejorar esas decisiones, (PS2) cuáles son sus desafíos y oportunidades y (PS3) cómo puede ayudar a los médicos a brindar una mejor atención a los pacientes.

2. METODOLOGÍA

2.1 Diseño

Se realizó una revisión sistemática con síntesis narrativa, de diseño descriptivo y analítico, en dos fases. La primera cubrió las publicaciones de 2019 a 2023. La segunda actualizó la búsqueda hasta el 12 de marzo de 2026. El proceso se reportó según la declaración PRISMA 2020 para revisiones actualizadas (Page et al., 2021) y la búsqueda según PRISMA-S (Rethlefsen et al., 2021). La revisión trabajó con publicaciones y no con personas, por lo que no requirió aprobación ética.

2.2 Criterios de selección

En ambas fases se buscaron estudios en inglés sobre la influencia de la IA en la toma de decisiones clínicas. La primera fase incluyó estudios cuantitativos observacionales y estudios cualitativos sobre las experiencias de los profesionales, publicados entre el 1 de enero de 2019 y el 31 de diciembre de 2023,

y excluyó revisiones sistemáticas, estudios no publicados, estudios en animales, duplicados, estudios con información insuficiente y estudios de baja calidad.

En la actualización, los criterios se aplicaron de forma más estricta. Se incluyeron estudios primarios cuantitativos (observacionales o experimentales, incluidos los ensayos aleatorizados), cualitativos o mixtos, publicados entre el 1 de enero de 2024 y el 12 de marzo de 2026, que evaluaran al menos uno de tres aspectos: profesionales de la salud que usan IA para decidir, con algún resultado sobre las decisiones, el desempeño, el trabajo o los pacientes; la comparación entre recomendaciones de la IA y decisiones de clínicos en casos reales o realistas; o las experiencias, la confianza o la intención de uso de profesionales en ejercicio. Se excluyeron las revisiones de cualquier tipo, los protocolos, los editoriales y artículos de opinión, los resúmenes de congresos, los preprints, los estudios que solo desarrollaban o validaban un modelo sin evaluar decisiones de clínicos y los estudios con población exclusivamente estudiantil, de pacientes o animal.

2.3 Fuentes y estrategia de búsqueda

La primera fase consultó Web of Science, Scopus, PubMed y Springer Link con cuatro bloques de términos aplicados principalmente al título. La actualización se hizo el 10 de julio de 2026 en el Directorio de Revistas de Acceso Abierto (DOAJ), mediante su interfaz de programación. Se eligió esta fuente porque permite búsquedas booleanas reproducibles y porque sus artículos tienen texto completo abierto, condición necesaria para valorar la calidad. Su límite de longitud obligó a concentrar la estrategia en los dos conceptos centrales, la IA y la decisión clínica, aplicados al título (Tabla 1). Los registros publicados después de marzo de 2026 se eliminaron antes del cribado. Además, se añadieron cinco registros identificados por otros métodos, en concreto mediante una búsqueda dirigida de estudios primarios recientes en revistas de alto impacto.

Tabla 1
Estrategias de búsqueda de las dos fases

Fase	Fuentes y periodo	Estrategia
Primera (versión previa)	Web of Science, Scopus, PubMed y Springer Link; 2019 a 2023	("artificial intelligence" OR AI OR "machine learning" OR "deep learning" OR "natural language processing" OR "computer vision") AND (ethics OR safety OR bias OR interoperability OR scalability OR "cost-effectiveness") AND ("clinical decision making" OR diagnosis OR prognosis OR treatment OR monitoring OR prevention) AND (medicine OR "medical specialty" OR disease OR symptom OR procedure)
Actualización	DOAJ; enero de 2024 a marzo de 2026.	bibjson.title:(AI OR "artificial intelligence" OR "machine learning" OR "deep learning") AND bibjson.title:"clinical decision" AND bibjson.year:[2024 TO 2026], con eliminación posterior de los registros publicados después del 12 de marzo de 2026
Otros métodos	Búsqueda dirigida en revistas biomédicas; 2024 a 2026	Ensayos y estudios primarios sobre IA y decisión clínica identificados en la literatura reciente

Nota. DOAJ: *Directory of Open Access Journals*. En la primera fase los términos de un mismo bloque se combinaron con OR y los bloques entre sí con AND.

2.4 Selección y valoración de la calidad

La selección siguió las fases de PRISMA: identificación, cribado por título, revisión del resumen, lectura del texto completo y valoración de la calidad. La calidad se valoró en ambas fases con diez preguntas derivadas de las pautas de Kitchenham y Charters (2007), con respuestas sí (1 punto), parcialmente (0,5) y no (0). Las preguntas valoraron la claridad de los objetivos, la fundamentación en estudios previos, la adecuación del método, la descripción de la recolección y el análisis de datos, la justificación del diseño, la correspondencia de los resultados con los objetivos, el uso de una metodología analítica, la respuesta a todas las preguntas, la discusión crítica y la coherencia de las conclusiones. Se exigió un mínimo de 7,5 sobre 10 (Castillo-Parra et al., 2022). En la actualización, el cribado, la valoración y la extracción se realizaron sobre el texto completo.

2.5 Extracción y síntesis

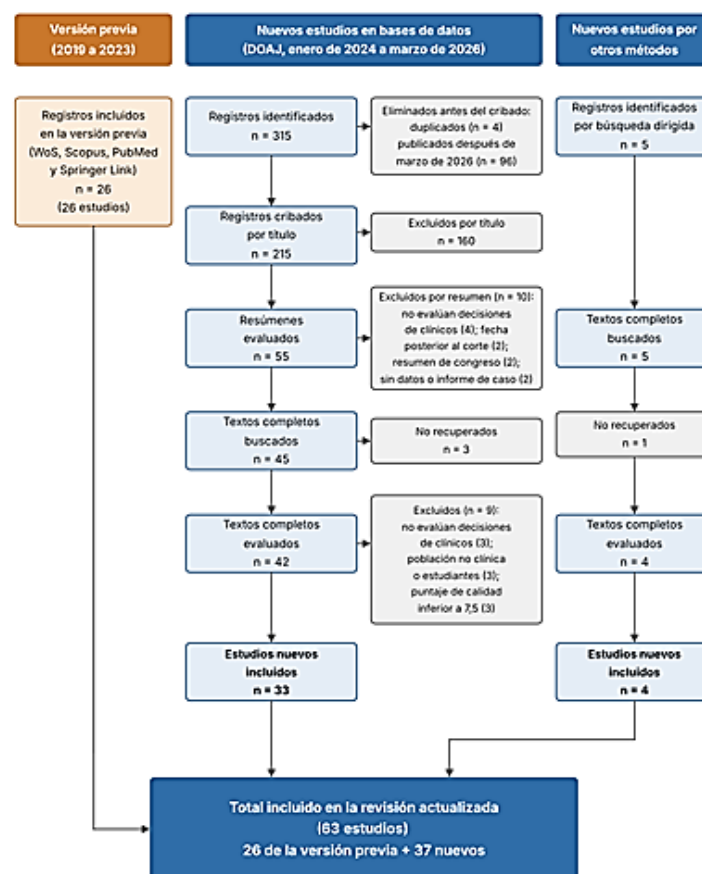
De cada estudio se extrajeron el país, el área clínica, el diseño, los participantes, el tipo de IA, la comparación y los hallazgos principales. Los diseños se agruparon en ensayos aleatorizados, estudios experimentales o de lectores, observacionales, comparativos con casos clínicos, encuestas, cualitativos, mixtos y Delphi. El tipo de IA se agrupó en predicción con datos clínicos, análisis de imagen y señales, modelos de lenguaje de gran tamaño e IA genérica o no especificada. La síntesis fue narrativa, se organizó por pregunta de investigación y comparó ambos periodos. No se realizó metaanálisis por la heterogeneidad de diseños y desenlaces, ni se graduó la certeza de la evidencia. Las revisiones, los metaanálisis y los documentos normativos se usaron solo para discutir los hallazgos y no se incluyeron como estudios.

3. RESULTADOS

3.1 Selección de los estudios

La versión previa aportó 26 registros que corresponden a 25 estudios, porque una referencia se registró dos veces (Ledziński y Grześk, 2023). Habían sido seleccionados de 799 registros, de los cuales 413 pasaron a lectura completa y 39 a la valoración de calidad. En la actualización, DOAJ devolvió 315 registros. Se eliminaron 4 duplicados y 96 publicados después de marzo de 2026, se cribaron 215 títulos y se revisaron 55 resúmenes. De los 45 textos completos buscados, 3 no pudieron recuperarse, 6 no cumplían los criterios y 3 no alcanzaron el puntaje mínimo, de modo que se incluyeron 33 estudios. Por otros métodos se identificaron 5 registros, uno no pudo recuperarse y 4 se incluyeron. En total, la revisión reúne 63 estudios, 26 de la primera fase y 37 de la actualización (Figura 1).

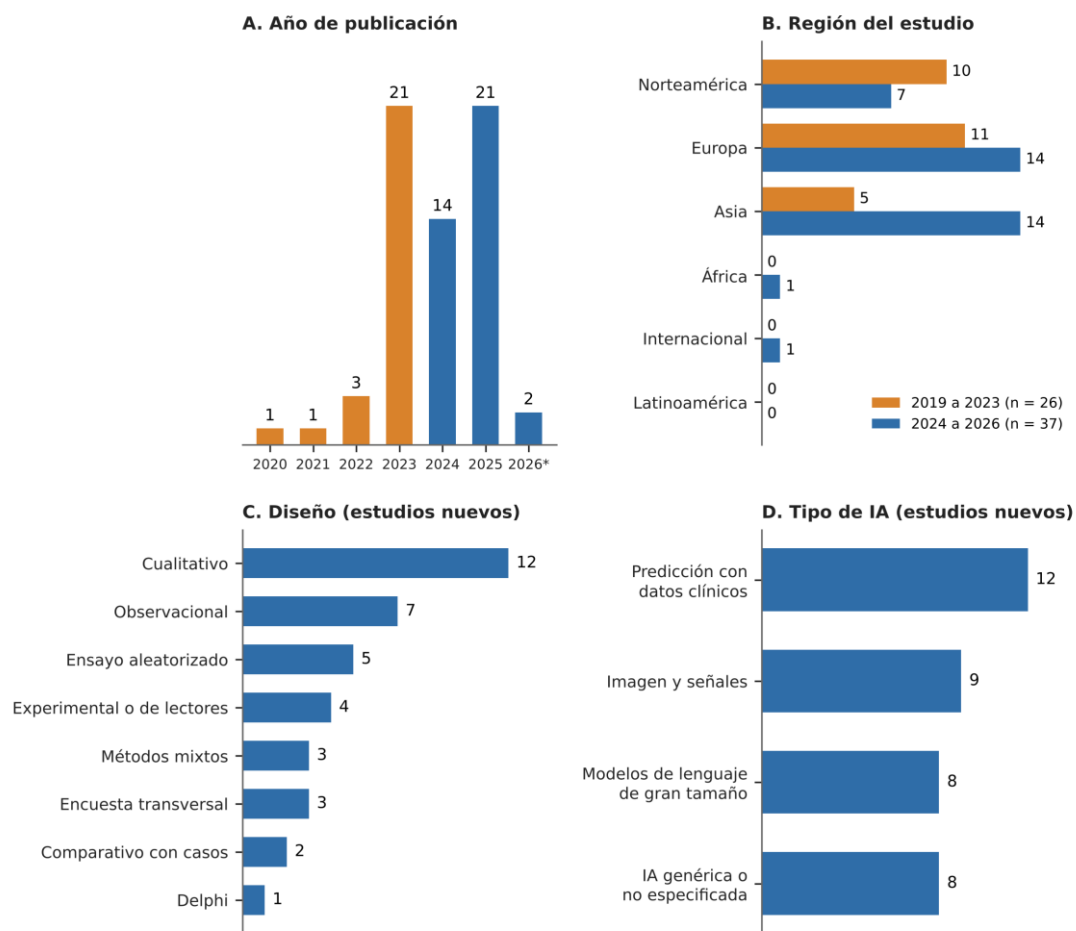
Figura 1
Diagrama de flujo PRISMA 2020 de la revisión actualizada



3.2 Características de los estudios

Los dos periodos son muy distintos (Tabla 2 y Figura 2). Los registros de 2019 a 2023 se concentraron en 2023 y procedieron de Norteamérica, Europa y Asia. Incluían revisiones, un protocolo, un informe de caso y preprints, y solo un estudio aportó cifras de desempeño (Juang et al., 2022). Los 37 estudios de la actualización son primarios: 12 cualitativos, 7 observacionales, 5 ensayos aleatorizados, 4 experimentales o de lectores, 3 mixtos, 3 encuestas, 2 comparativos con casos clínicos y 1 Delphi. Europa y Asia aportaron 14 cada una, Norteamérica 7, África 1 y hubo un estudio internacional. Ninguno de los dos periodos incluyó estudios de Latinoamérica. Por tipo de IA, 12 estudios evaluaron modelos de predicción con datos clínicos, 9 herramientas de imagen y señales, 8 modelos de lenguaje de gran tamaño y 8 sistemas genéricos o no especificados.

Figura 2
Distribución de los estudios incluidos según año, región, diseño y tipo de IA



* Enero a marzo de 2026. Naranja: versión previa de la revisión; azul: actualización.

Tabla 2
Características de los estudios incluidos en cada periodo

Característica	2019 a 2023 (26 estudios)	2024 a marzo de 2026 (37 estudios)
Tipo de documento	17 declarados cualitativos y 9 cuantitativos; incluye revisiones, un protocolo, un informe de caso y preprints	Solo estudios primarios; 5 ensayos aleatorizados y 4 experimentales o de lectores
Región	Norteamérica 10, Europa 11, Asia 5	Europa 14, Asia 14, Norteamérica 7, África 1, internacional 1
Latinoamérica	0	0
Modelos de lenguaje	1	8

Característica	2019 a 2023 (26 estudios)	2024 a marzo de 2026 (37 estudios)
Datos cuantitativos de desempeño o decisión	1 estudio	18 estudios, 4 con desenlaces del paciente o del proceso asistencial
Puntaje de calidad, media (rango)	8,85 (8,0 a 10,0)	8,88 (7,5 a 10,0)

3.3 Calidad de los estudios

En la actualización los puntajes fueron de 7,5 a 10,0, con una media de 8,88, y tres estudios quedaron fuera por no alcanzar el corte. Los criterios más débiles fueron la adecuación del método (media de 0,77 sobre 1), la justificación del diseño (0,80) y el uso de una metodología analítica explícita (0,82). En la primera fase, la media fue de 8,85 y el criterio más débil fue también el análisis de los resultados (0,77). Como casi todos los estudios puntuaron 8,0 o más, el instrumento distinguió poco entre ellos. Se puede evidenciar que todos los estudios incluidos superaron el umbral.

3.4 Utilidad de la IA para mejorar la toma de decisiones clínicas (PS1)

En la primera fase, la utilidad se describía sobre todo como posibilidad: apoyo a la decisión quirúrgica (Loftus et al., 2020), a la predicción y el tratamiento de la sepsis (Wu et al., 2021), al diagnóstico de enfermedades genéticas (Mahwish et al., 2022) y a la planificación de la radioterapia (Hernandez et al., 2023). La actualización aporta datos medidos, y estos muestran beneficios claros en tareas bien delimitadas.

En un ensayo con 105 934 mujeres, el tamizaje mamográfico con apoyo de IA detectó 6,4 cánceres por 1000 frente a 5,0, sin aumentar los falsos positivos, y redujo la carga de lectura en 44,2 % (Hernström et al., 2025). En otro ensayo, los residentes apoyados por un modelo identificaron correctamente el 95,3 % de los sonidos de auscultación frente al 89,3 % del grupo control (Hirosawa et al., 2024). En traumatismos de muñeca, la IA permitió diagnosticar sin ayuda adicional con más frecuencia (75 % frente a 52 %) y redujo las tomografías solicitadas (14 % frente a 28 %), aunque la exactitud global no cambió (Keller et al., 2024). En un hospital de Taiwán, un sistema que predice la resistencia antibiótica se asoció con menor mortalidad a 14 días (11,5 % frente a 15,1 %), si bien la adecuación de la prescripción se midió por autoinforme (T.-H. Lin et al., 2025).

Los estudios de implementación muestran efectos sobre el proceso asistencial, es así que en urgencias, un sistema que predice intervenciones críticas redujo el tiempo hasta varias de ellas, como la vía arterial y la oxigenoterapia (Chang et al., 2025). En reproducción asistida, los médicos que usaron una herramienta de IA indicaron dosis menores de hormona folículo estimulante sin perder ovocitos maduros (Bixby y Miller, 2025). En Kenia, el uso mensual de un asistente basado en un modelo de lenguaje integrado en la historia clínica pasó del 4 % al 47 % de las consultas en ocho meses (Obong'o et al., 2026). Otros estudios midieron la concordancia entre la IA y los clínicos, que fue alta en oncología mamaria (80,4 %, y 92,4 % tras revisar los casos discrepantes) (Xu et al., 2024), en perfiles pronósticos de dolor crónico (82,4 %) (Zmudzki et al., 2025) y baja en el ajuste de bombas de insulina (Sorokin et al., 2024).

Sin embargo, la disponibilidad de herramientas de inteligencia artificial no se traduce necesariamente en mejores decisiones clínicas. Un ensayo reciente mostró que el acceso de los médicos a un modelo de lenguaje no produjo mejoras significativas en el razonamiento diagnóstico (76 % frente a 74 %), pese a que el modelo, evaluado de manera aislada, alcanzó una precisión diagnóstica del 92 % (Goh et al., 2024). Este hallazgo pone de manifiesto que el beneficio potencial de estas tecnologías no depende exclusivamente de su rendimiento técnico, sino también de su adecuada integración en los flujos de trabajo clínicos y de la capacidad de los profesionales para interpretar y utilizar críticamente sus resultados. En otro experimento del mismo grupo, la exactitud de los médicos subió cerca de 18 puntos tras ver las recomendaciones del modelo, sin ampliar las diferencias por sexo o raza del paciente (Goh et al., 2025). Con modelos de lenguaje en casos reales de dolor abdominal, los médicos acertaron entre

el 87,5 % y el 92,5 % y los modelos entre el 58,8 % y el 67,8 % (Hager et al., 2024). En radioterapia del melanoma uveal, un modelo superó a los oncólogos y otros dos quedaron por debajo (Wang y Wang, 2025), y en planificación de implantes dentales la concordancia entre el modelo y los clínicos fue baja (Gönültaş et al., 2025). La forma en que los médicos introducen la información en el modelo no se asoció con mejores resultados (Siden et al., 2025).

3.5 Desafíos y oportunidades (PS2)

Los desafíos de la primera fase, como la calidad y el sesgo de los datos, la interpretabilidad, la responsabilidad del médico y la regulación (Kulkarni y Singh, 2023; Benzinger et al., 2023; Lawton et al., 2024; Ledziński y Grześk, 2023), reaparecen en la actualización, ahora con evidencia empírica.

El primer desafío es la dependencia mal calibrada. Entre 223 dermatólogos, la IA elevó la exactitud solo de 63 % a 65 %, porque los médicos conservaron sus aciertos pero rara vez adoptaron el consejo correcto del sistema (Küper et al., 2025). En la lectura de tomografías cerebrales, la IA aumentó la sensibilidad de los profesionales inexpertos de 59,3 % a 72,7 % pero redujo su especificidad de 65,5 % a 53,7 %, y los llevó a cambiar más decisiones sobre el destino del paciente (Choi et al., 2024). El segundo es la pérdida de habilidades: tras introducir la IA en cuatro centros, la detección de adenomas en colonoscopias sin IA bajó de 28,4 % a 22,4 % (Budzyń et al., 2025). El tercero es la fiabilidad de los modelos de lenguaje, que no siguieron las guías clínicas (Hager et al., 2024), dieron respuestas inconsistentes (Gönültaş et al., 2025) o no se ajustaron a las guías locales (Obong'o et al., 2026).

Los estudios cualitativos explican por qué los clínicos confían o no en la IA, y en este sentido necesitan conocer la población con que se entrenó el modelo y su desempeño en términos clínicos (Mingels et al., 2026), y aceptan la IA si es precisa, personalizada, transparente y mejora resultados (Dlugatch et al., 2024). Algunos malinterpretan qué datos usa el sistema (Shulha et al., 2024). La adopción depende de factores técnicos, organizativos y regulatorios (Schaffernak et al., 2025; Düvel et al., 2025), y de la integración con la historia clínica (Uihlein et al., 2024; Cooper et al., 2025). La responsabilidad sigue sin resolverse. El 73 % de los médicos entrevistados en Singapur consideró que la responsabilidad legal recae en quien toma la decisión final (Huang et al., 2024), y otros clínicos anticiparon angustia moral cuando la recomendación choque con su juicio o con la falta de recursos (Whitney et al., 2025). También preocupa la equidad en la asignación de recursos (Elgin y Elgin, 2024).

Las encuestas muestran que la intención de uso depende sobre todo de la utilidad esperada y de la disposición personal a innovar (Yu et al., 2025; Zhao et al., 2025). En farmacia hospitalaria, la utilidad percibida fue el principal predictor y el 82,6 % nunca había usado un sistema de apoyo (Ghozali et al., 2025). En medicina de desastres, los expertos solo alcanzaron consenso para usos logísticos y no para decisiones clínicas críticas (Franc et al., 2025). Entre las oportunidades destacan la segunda opinión en la prescripción compleja (d'Elia et al., 2025), el apoyo a la decisión compartida (Gunlicks-Stoessel et al., 2024) y la formación de residentes (Whitney et al., 2025). En la primera fase ya se proponía ampliar la decisión compartida con el paciente (Lorenzini et al., 2023).

3.6 Aporte de la IA a la atención del paciente (PS3)

En la primera fase, el aporte a la atención se describía como potencial, ya sea en la detección más temprana, personalización y manejo de más información (Rathore y Rathore, 2023; Wehkamp et al., 2023; Mehta, 2023). En la actualización, cuatro estudios midieron efectos sobre el paciente o el proceso asistencial: más cánceres detectados (Hernström et al., 2025), menor mortalidad (T.-H. Lin et al., 2025), dosis hormonales menores con igual eficacia (Bixby y Miller, 2025) e intervenciones críticas más rápidas (Chang et al., 2025).

La IA también puede favorecer la decisión compartida. En dolor crónico, el perfil pronóstico se discutió con el paciente en el 47,1 % de los casos (Zmudzki et al., 2025), y en depresión adolescente los terapeutas usaron las recomendaciones para decidir con las familias (Gunlicks-Stoessel et al., 2024). Sin embargo, los 20 profesionales de un estudio en multimorbilidad expresaron preocupación por la relación con el paciente y la mayoría exigió que la decisión final siga siendo humana (Cooper et al., 2025). Las

enfermeras neonatales la ven como apoyo y no como sustituto del juicio clínico (Alruwaili et al., 2025). La Tabla 3 resume la síntesis por pregunta y periodo.

Tabla 3
Síntesis por pregunta de investigación y periodo

Pregunta	Estudios 2019 a 2023	Estudios 2024 a 2026	Síntesis
PS1. Utilidad	Loftus et al. (2020); Wu et al. (2021); Damato et al. (2022); Juang et al. (2022); Mahwish et al. (2022); Benzinger et al. (2023); Brown et al. (2023); Ledziński y Grześk (2023); Hernandez et al. (2023)	Alruwaili et al. (2025); Bixby y Miller (2025); Budzyń et al. (2025); Chang et al. (2025); Choi et al. (2024); Düvel et al. (2025); Goh et al. (2024, 2025); Gunlicks-Stoessel et al. (2024); Guo et al. (2024); Gönüldaş et al. (2025); Hager et al. (2024); Hernström et al. (2025); Hirosawa et al. (2024); Keller et al. (2024); Küper et al. (2025); Obong'o et al. (2026); Siden et al. (2025); Sorokin et al. (2024); T.-H. Lin et al. (2025); Wang y Wang (2025); Xu et al. (2024); Zmudzki et al. (2025)	De la utilidad descrita a beneficios medidos en tamizaje, auscultación, prescripción y urgencias; resultados nulos o desfavorables con modelos de lenguaje cuando el médico no aprovecha la herramienta.
PS2. Desafíos y oportunidad es	Blagojević y Geroski (2023); Ambags et al. (2023); Bajgain et al. (2023); Kulkarni y Singh (2023); Lawton et al. (2024); Lorenzini et al. (2023); L. Wang et al. (2023); Yang et al. (2023); Almotairi et al. (2023); Ramgopal et al. (2023)	Alruwaili et al. (2025); Budzyń et al. (2025); Chang et al. (2025); Choi et al. (2024); Cooper et al. (2025); d'Elia et al. (2025); Dlugatch et al. (2024); Düvel et al. (2025); Elgin y Elgin (2024); Franc et al. (2025); Ghozali et al. (2025); Goh et al. (2024, 2025); Gunlicks-Stoessel et al. (2024); Gönüldaş et al. (2025); Hager et al. (2024); Huang et al. (2024); Küper et al. (2025); Mingels et al. (2026); Obong'o et al. (2026); Schaffernak et al. (2025); Shulha et al. (2024); Siden et al. (2025); Uihlein et al. (2024); Wang y Wang (2025); Whitney et al. (2025); Yu et al. (2025); Zhao et al. (2025)	Dependencia mal calibrada, falsos positivos, pérdida de habilidades, fiabilidad de los modelos de lenguaje, necesidad de transparencia, integración y responsabilidad clara. Oportunidades en la decisión compartida y la formación.
PS3. Atención al paciente	Iftimia et al. (2023); Rathore y Rathore (2023); Briganti (2023); Wehkamp et al. (2023); Mehta (2023); Santoro (2023); Benzinger et al. (2023)	Alruwaili et al. (2025); Bixby y Miller (2025); Chang et al. (2025); Cooper et al. (2025); d'Elia et al. (2025); Goh et al. (2025); Gunlicks-Stoessel et al. (2024); Hernström et al. (2025); Hirosawa et al. (2024); Keller et al. (2024); Mingels et al. (2026); Obong'o et al. (2026); Schaffernak et al. (2025); Shulha et al. (2024); T.-H. Lin et al. (2025); Zmudzki et al. (2025)	Cuatro estudios con efectos sobre el paciente o el proceso asistencial; apoyo a la decisión compartida, con preocupación por la relación clínico paciente.

Nota. Un estudio puede responder a más de una pregunta.

4. DISCUSIÓN

La respuesta a la pregunta principal es que la influencia de la IA en la toma de decisiones clínicas pasó, entre 2019 y 2026, de la promesa a la evidencia medida. Hasta 2023 la literatura describía posibilidades. Desde 2024 aparecen ensayos aleatorizados, estudios de implementación y estudios de lectores que muestran beneficios concretos, como más cánceres detectados con menos carga de trabajo (Hernström et al., 2025), mejor auscultación (Hirosawa et al., 2024) y decisiones más rápidas en urgencias (Chang et al., 2025). Otro ensayo, que no pudo incluirse por falta de acceso al texto completo, informó menor mortalidad con alertas electrocardiográficas basadas en IA (C.-S. Lin et al., 2024).

El hallazgo más importante es que el beneficio depende de la interacción entre el médico y la herramienta y no solo de la exactitud del modelo. El mismo tipo de modelo de lenguaje mejoró la exactitud de los médicos en un experimento (Goh et al., 2025) y no cambió su razonamiento en un ensayo, aunque por sí solo superaba a los médicos (Goh et al., 2024). Los dermatólogos rara vez aceptaron el consejo correcto de la IA (Küper et al., 2025), mientras que los profesionales inexpertos aceptaron demasiado sus falsos positivos (Choi et al., 2024). Ambas situaciones son formas de dependencia mal calibrada, ya descrita como sesgo de automatización en mamografía (Dratsch et al.,

2023). Por eso la formación en el uso de estas herramientas (Siden et al., 2025) y el diseño de la interacción importan tanto como el algoritmo.

La seguridad exige además validar cada herramienta en su contexto. Un modelo comercial de sepsis, evaluado en 38 455 hospitalizaciones, no detectó el 67 % de los casos (Wong et al., 2021), y los modelos de lenguaje evaluados en casos reales no siguieron las guías (Hager et al., 2024). La caída en la detección de adenomas tras el uso habitual de la IA (Budzyń et al., 2025) añade un riesgo que la primera fase solo intuía (Yang et al., 2023; Iftimia et al., 2023), el de que el médico pierda la capacidad de detectar los errores del sistema. Las revisiones recientes coinciden en que los ensayos favorables son frecuentes, pero pocos tienen bajo riesgo de sesgo o miden desenlaces en pacientes (Han et al., 2024; Wilhelm et al., 2025), y por eso guías como DECIDE-AI y TRIPOD+AI piden informar cómo y en qué población se evaluó cada sistema (Vasey et al., 2022; Collins et al., 2024).

La confianza y la responsabilidad siguen sin resolverse. Los clínicos piden saber con qué datos se entrenó el modelo y cómo rinde en términos clínicos (Mingels et al., 2026), y condicionan su confianza a la evidencia de mejores resultados (Dlugatch et al., 2024). Estas condiciones coinciden con una revisión de 27 estudios que identificó la transparencia, la formación y la fiabilidad clínica como factores de confianza (Tun et al., 2025). La advertencia sobre el médico como “sumidero de responsabilidad” (Lawton et al., 2024) se confirma en la percepción de los médicos de que la responsabilidad legal recae en quien decide (Huang et al., 2024) y en la angustia moral que anticipan (Whitney et al., 2025). Estos hallazgos se relacionan con los principios de rendición de cuentas de la OMS (2021) y con la supervisión humana que exige el reglamento europeo (Unión Europea, 2024).

La ausencia de estudios latinoamericanos en los dos periodos es un hallazgo en sí mismo. El estudio de Kenia muestra que un asistente basado en un modelo de lenguaje puede adoptarse en atención primaria de un país de ingresos medianos bajos (Obong'o et al., 2026), pero no hay datos equivalentes para la región. Las condiciones tampoco son favorables, ya que solo 10 países de América Latina y el Caribe tienen marcos regulatorios avanzados para la historia clínica electrónica (Farias et al., 2023), y la mayoría de profesionales de la región usa herramientas de IA sin formación formal (OPS, 2025). En Ecuador, el carácter sensible de los datos de salud (Asamblea Nacional del Ecuador, 2021) obliga a que cualquier sistema entrenado o evaluado con datos locales cumpla requisitos estrictos de protección. Esta revisión no permite saber si la ausencia refleja falta de investigación o el efecto de buscar solo en inglés.

Para la práctica, los resultados sugieren cuatro pautas: usar la IA en tareas bien delimitadas y validadas localmente, formar a los médicos para usarla y cuestionarla, vigilar la dependencia excesiva o insuficiente y la pérdida de habilidades, y mantener la decisión compartida con el paciente.

5. CONCLUSIONES

Entre 2019 y 2026, la evidencia sobre la IA en la toma de decisiones clínicas pasó de la descripción de posibilidades a estudios que miden su efecto. La IA mejora decisiones concretas, como el tamizaje mamográfico, la auscultación, la prescripción antibiótica y la respuesta en urgencias, cuando se aplica a tareas bien delimitadas y validadas.

El beneficio no depende solo de la exactitud del modelo, sino de cómo lo usa el médico. La dependencia excesiva o insuficiente, los falsos positivos, la pérdida de habilidades y la falta de fiabilidad de los modelos de lenguaje en casos reales son los principales riesgos. Los profesionales aceptan la IA como apoyo y no como sustituto, y piden transparencia, integración con la historia clínica y claridad sobre la responsabilidad.

Las próximas investigaciones deberían realizar ensayos con desenlaces en pacientes, seguir las guías de reporte específicas para IA e incluir poblaciones latinoamericanas, con atención a la protección de los datos de salud y a la formación de los profesionales.

Declaraciones

- **Financiación:** Los autores declaran que no existió financiación para el desarrollo de la investigación.
- **Consideraciones éticas:** Al tratarse de una revisión sistemática de literatura basada en fuentes secundarias publicadas, no requirió consentimiento informado ni aprobación ética.
- **Declaración de conflicto de interés:** Los autores no tiene ningún conflicto de interés relacionado con la investigación, los resultados o la publicación de este artículo.
- **Contribución de los autores al trabajo:** WJZB: Idea, Revisión de literatura (estado del arte; Metodología; Búsqueda, selección y NMMA extracción de datos (Análisis de datos; Resultados; Discusión y conclusiones; Redacción.
- **Declaración de uso de inteligencia artificial:** Los autores declaran no haber hecho uso de la inteligencia artificial para la elaboración de este proyecto y aportación académica.

Referencias

- Almotairi, K. H., Hussein, A. M. A., Abualigah, L., Abujayyab, S. K. M., Mahmoud, E. H., Ghanem, B. O., y Gandomi, A. H. (2023). Impact of artificial intelligence on COVID-19 pandemic: A survey of image processing, tracking of disease, prediction of outcomes, and computational medicine. *Big Data and Cognitive Computing*, 7(1), 11. <https://doi.org/10.3390/bdcc7010011>
- Alruwaili, A. N., Alshammari, A. M., Alhaiti, A., Elsharkawy, N. B., Ali, S. I., y Ramadan, O. M. E. (2025). Neonatal nurses' experiences with generative AI in clinical decision-making: A qualitative exploration in high-risk NICUs. *BMC Nursing*, 24, 386. <https://doi.org/10.1186/s12912-025-03044-6>
- Ambags, E. L., Capitoli, G., L'Imperio, V., Provenzano, M., Nobile, M. S., y Liò, P. (2023). *Assisting clinical practice with fuzzy probabilistic decision trees* [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2304.07788>
- Asamblea Nacional del Ecuador. (2021). Ley Orgánica de Protección de Datos Personales. *Registro Oficial*, Quinto Suplemento 459, 26 de mayo de 2021.
- Bajgain, B., Lorenzetti, D., Lee, J., y Sauro, K. (2023). Determinants of implementing artificial intelligence-based clinical decision support tools in healthcare: A scoping review protocol. *BMJ Open*, 13(2), e068373. <https://doi.org/10.1136/bmjopen-2022-068373>
- Benzinger, L., Ursin, F., Balke, W. T., Kacprowski, T., y Salloch, S. (2023). Should artificial intelligence be used to support clinical ethical decision-making? A systematic review of reasons. *BMC Medical Ethics*, 24(1). <https://doi.org/10.1186/s12910-023-00929-6>
- Bixby, C. J., y Miller, B. (2025). Real-world use of an artificial intelligence-powered clinical decision support tool for ovarian stimulation. *F&S Reports*, 6(2), 140-146. <https://doi.org/10.1016/j.xfre.2025.01.015>
- Blagojević, A., y Geroski, T. (2023). A review of the application of artificial intelligence in medicine: From data to personalised models. En N. Filipović (Ed.), *Applied artificial intelligence: Medicine, biology, chemistry, financial, games, engineering* (pp. 271-305). Springer. https://doi.org/10.1007/978-3-031-29717-5_17
- Briganti, G. (2023). Artificial intelligence in psychiatry. *Psychiatria Danubina*, 35(Supl. 2), 15-19.
- Brown, C., Nazeer, R., Gibbs, A., Le Page, P., y Mitchell, A. R. (2023). Breaking bias: The role of artificial intelligence in improving clinical decision-making. *Cureus*, 15(3), e36415. <https://doi.org/10.7759/cureus.36415>
- Budzyń, K., Romańczyk, M., Kitala, D., Kołodziej, P., Bugajski, M., Adami, H. O., Blom, J., Buszkiewicz, M., Halvorsen, N., Hassan, C., Romańczyk, T., Holme, Ø., Jarus, K., Fielding, S., Kunar, M., Pellise, M., Pilonis, N., Kamiński, M. F., Kalager, M., ... Mori, Y. (2025). Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: A multicentre, observational study. *The Lancet Gastroenterology & Hepatology*, 10(10), 896-903. [https://doi.org/10.1016/S2468-1253\(25\)00133-5](https://doi.org/10.1016/S2468-1253(25)00133-5)
- Castillo-Parra, B., Hidalgo-Cajo, B. G., Váscquez-Barrera, M., y Oleas-López, J. (2022). Gamification in higher education: A review of the literature. *World Journal on Educational Technology: Current Issues*, 14(3), 797-816. <https://doi.org/10.18844/wjet.v14i3.7341>
- Chang, H., Yu, J. Y., Park, H., Song, Y. J., Heo, S., Park, J. E., Lee, G. T., Lee, S. U., Kim, T., Yoon, H., Hwang, S. Y., y Cha, W. C. (2025). From prediction to action: A retrospective observational study on the real-world implementation of Critical Interventions (CrIs), an AI-based clinical decision support system changing clinical behavior in the emergency department. *BMC Medical Informatics and Decision Making*, 25, 405. <https://doi.org/10.1186/s12911-025-03234-x>

- Choi, S. Y., Kim, J. H., Chung, H. S., Lim, S., Kim, E. H., y Choi, A. (2024). Impact of a deep learning-based brain CT interpretation algorithm on clinical decision-making for intracranial hemorrhage in the emergency department. *Scientific Reports*, 14, 22292. <https://doi.org/10.1038/s41598-024-73589-0>
- Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., van Smeden, M., Boulesteix, A.-L., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., ... Logullo, P. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*, 385, e078378. <https://doi.org/10.1136/bmj-2023-078378>
- Cooper, J., Haroon, S., Crowe, F., Nirantharakumar, K., Jackson, T., Fitzsimmons, L., Hathaway, E., Flanagan, S., Marshall, T., Jackson, L. J., Gunathilaka, N., D'Elia, A., Morris, S. G., y Greenfield, S. (2025). Perspectives of health care professionals on the use of AI to support clinical decision-making in the management of multiple long-term conditions: Interview study. *Journal of Medical Internet Research*, 27, e71980. <https://doi.org/10.2196/71980>
- Correia, E. A. S., y Valério Netto, A. (2025). Suporte à decisão clínica baseado em IA: Revisão de escopo. *Journal of Health Informatics*, 17(1). <https://doi.org/10.59681/2175-4411.v17.2025.1457>
- Damato, V., Campiglia, P., y Bassano, C. (2022). Artificial intelligence algorithms in precision medicine: A new approach in clinical decision-making. En *The human side of service engineering* (AHFE International Conference). AHFE Open Access. <https://doi.org/10.54941/ahfe1002561>
- d'Elia, A., Morris, S. G., Cooper, J., Nirantharakumar, K., Jackson, T., Marshall, T., Fitzsimmons, L., Jackson, L. J., Crowe, F., Haroon, S., Greenfield, S., y Hathaway, E. (2025). Perceptions of an AI-based clinical decision support tool for prescribing in multiple long-term conditions: A qualitative study of general practice clinicians in England. *BMJ Open*, 15(11), e102833. <https://doi.org/10.1136/bmjopen-2025-102833>
- Dlugatch, R., Georgieva, A., y Kerasidou, A. (2024). AI-driven decision support systems and epistemic reliance: A qualitative study on obstetricians' and midwives' perspectives on integrating AI-driven CTG into clinical decision making. *BMC Medical Ethics*, 25, 6. <https://doi.org/10.1186/s12910-023-00990-1>
- Dratsch, T., Chen, X., Rezazade Mehrizi, M., Kloeckner, R., Mähringer-Kunz, A., Püsken, M., Baeßler, B., Sauer, S., Maintz, D., y Pinto dos Santos, D. (2023). Automation bias in mammography: The impact of artificial intelligence BI-RADS suggestions on reader performance. *Radiology*, 307(4), e222176. <https://doi.org/10.1148/radiol.222176>
- Düvel, J. A., Lampe, D., Kirchner, M., Elkenkamp, S., Cimiano, P., Düsing, C., Marchi, H., Schmiegel, S., Fuchs, C., Claßen, S., Meier, K.-L., Borgstedt, R., Rehberg, S., y Greiner, W. (2025). An AI-based clinical decision support system for antibiotic therapy in sepsis (KINBIOTICS): Use case analysis. *JMIR Human Factors*, 12, e66699. <https://doi.org/10.2196/66699>
- Elgin, C. Y., y Elgin, C. (2024). Ethical implications of AI-driven clinical decision support systems on healthcare resource allocation: A qualitative study of healthcare professionals' perspectives. *BMC Medical Ethics*, 25, 148. <https://doi.org/10.1186/s12910-024-01151-8>
- Farias, M. A., Badino, M., Marti, M., Báscolo, E., García Saisó, S., y D'Agostino, M. (2023). La transformación digital como estrategia para el fortalecimiento de las funciones esenciales de salud pública en las Américas. *Revista Panamericana de Salud Pública*, 47, e150. <https://doi.org/10.26633/RPSP.2023.150>
- Franc, J. M., Verde, M., Bonney, J., Hung, K. K. C., Cuthbertson, J., Raffee, L., Serra, E., y Caviglia, M. (2025). Applications of artificial intelligence-guided clinical decision support in disaster medicine: An international delphi study. *Frontiers in Disaster and Emergency Medicine*, 3, 1698372. <https://doi.org/10.3389/femer.2025.1698372>
- Ghozali, M. T., Satibi, y Fortwengel, G. (2025). Adoption intentions toward AI-based clinical decision support tools: A TAM study on hospital pharmacists. *Journal of Information and Communication Technology (Journal of ICT)*, 24(3), 44-69. <https://doi.org/10.32890/jict2025.24.3.3>
- Goh, E., Bunning, B., Khoong, E. C., Gallo, R. J., Milstein, A., Centola, D., y Chen, J. H. (2025). Physician clinical decision modification and bias assessment in a randomized controlled trial of AI assistance. *Communications Medicine*, 5(1), 59. <https://doi.org/10.1038/s43856-025-00781-2>
- Goh, E., Gallo, R., Hom, J., Strong, E., Weng, Y., Kerman, H., Cool, J. A., Kanjee, Z., Parsons, A. S., Ahuja, N., Horvitz, E., Yang, D., Milstein, A., Olson, A. P. J., Rodman, A., y Chen, J. H. (2024). Large language model influence on diagnostic reasoning: A randomized clinical trial. *JAMA Network Open*, 7(10), e2440969. <https://doi.org/10.1001/jamanetworkopen.2024.40969>

- Gomez-Cabello, C. A., Borna, S., Pressman, S., Haider, S. A., Haider, C. R., y Forte, A. J. (2024). Artificial-intelligence-based clinical decision support systems in primary care: A scoping review of current clinical implementations. *European Journal of Investigation in Health, Psychology and Education*, 14(3), 685-698. <https://doi.org/10.3390/ejihpe14030045>
- Gönüldaş, F., Bulut, E. A., Özbey, Ö. N., y Öztürk, C. (2025). Clinical decision-making in implant planning: A comparison between human and artificial intelligence. *Applied Sciences*, 15(19), 10744. <https://doi.org/10.3390/app151910744>
- Gunlicks-Stoessel, M., Liu, Y., Parkhill, C., Morrell, N., Choy-Brown, M., Mehus, C., Hetler, J., y August, G. (2024). Adolescent, parent, and provider attitudes toward a machine learning based clinical decision support system for selecting treatment for youth depression. *BMC Medical Informatics and Decision Making*, 24(1), 4. <https://doi.org/10.1186/s12911-023-02410-1>
- Guo, Y., Huang, C., Sheng, Y., Zhang, W., Ye, X., Lian, H., Xu, J., y Chen, Y. (2024). Improve the efficiency and accuracy of ophthalmologists' clinical decision-making based on AI technology. *BMC Medical Informatics and Decision Making*, 24, 192. <https://doi.org/10.1186/s12911-024-02587-z>
- Hager, P., Jungmann, F., Holland, R., Bhagat, K., Hubrecht, I., Knauer, M., Vielhauer, J., Makowski, M., Braren, R., Kaissis, G., y Rueckert, D. (2024). Evaluation and mitigation of the limitations of large language models in clinical decision-making. *Nature Medicine*, 30(9), 2613-2622. <https://doi.org/10.1038/s41591-024-03097-1>
- Han, R., Acosta, J. N., Shakeri, Z., Ioannidis, J. P. A., Topol, E. J., y Rajpurkar, P. (2024). Randomised controlled trials evaluating artificial intelligence in clinical practice: A scoping review. *The Lancet Digital Health*, 6(5), e367-e373. [https://doi.org/10.1016/S2589-7500\(24\)00047-5](https://doi.org/10.1016/S2589-7500(24)00047-5)
- Hernandez, S., Nguyen, C., Parkes, J., Burger, H., Rhee, D. J., Netherton, T., Mumme, R., Vega, J. G. D. L., Duryea, J., Leone, A., Paulino, A. C., Cardenas, C., Howell, R., Fuentes, D., Pollard-Larkin, J., y Court, L. (2023). Automating the treatment planning process for 3D-conformal pediatric craniospinal irradiation therapy. *Pediatric Blood & Cancer*, 70(3), e30164. <https://doi.org/10.1002/pbc.30164>
- Hernström, V., Josefsson, V., Sartor, H., Schmidt, D., Larsson, A.-M., Hofvind, S., Andersson, I., Rosso, A., Hagberg, O., y Lång, K. (2025). Screening performance and characteristics of breast cancer detected in the Mammography Screening with Artificial Intelligence trial (MASAI): A randomised, controlled, parallel-group, non-inferiority, single-blinded, screening accuracy study. *The Lancet Digital Health*, 7(3), e175-e183. [https://doi.org/10.1016/S2589-7500\(24\)00267-X](https://doi.org/10.1016/S2589-7500(24)00267-X)
- Hirosawa, T., Sakamoto, T., Harada, Y., Tokumasu, K., y Shimizu, T. (2024). Clinical decision support system using a machine learning model to assist simultaneous cardiopulmonary auscultation: Open-label randomized controlled trial. *Digital Health*, 10, 20552076241233689. <https://doi.org/10.1177/20552076241233689>
- Huang, Z., Lim, H. Y.-F., Ow, J. T., Sun, S. H.-L., y Chow, A. (2024). Doctors' perception on the ethical use of AI-enabled clinical decision support systems for antibiotic prescribing recommendations in Singapore. *Frontiers in Public Health*, 12, 1420032. <https://doi.org/10.3389/fpubh.2024.1420032>
- Ifitimia, N., Pandya, R., y Mahoney, F. (2023). *New advances in artificial intelligence for biomedical research and clinical decision-making* [Preprint]. Preprints.org. <https://doi.org/10.20944/preprints202306.0243.v1>
- Juang, W. C., Hsu, M. H., Cai, Z. X., y Chen, C. M. (2022). Developing an AI-assisted clinical decision support system to enhance in-patient holistic health care. *PLOS ONE*, 17(10), e0276501. <https://doi.org/10.1371/journal.pone.0276501>
- Keller, M., Rohner, M., y Honigmann, P. (2024). The potential benefit of artificial intelligence regarding clinical decision-making in the treatment of wrist trauma patients. *Journal of Orthopaedic Surgery and Research*, 19, 579. <https://doi.org/10.1186/s13018-024-05063-6>
- Kitchenham, B., y Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Informe técnico EBSE-2007-01). Keele University y University of Durham.
- Kulkarni, P. A., y Singh, H. (2023). Artificial intelligence in clinical diagnosis: Opportunities, challenges, and hype. *JAMA*, 330(4), 317-318. <https://doi.org/10.1001/jama.2023.11440>
- Küper, A., Lodde, G. C., Livingstone, E., Schadendorf, D., y Krämer, N. (2025). Psychological factors influencing appropriate reliance on AI-enabled clinical decision support systems: Experimental web-based study among dermatologists. *Journal of Medical Internet Research*, 27, e58660. <https://doi.org/10.2196/58660>

- Lawton, T., Morgan, P., Porter, Z., Hickey, S., Cunningham, A., Hughes, N., Iacovides, I., Jia, Y., Sharma, V., y Habli, I. (2024). Clinicians risk becoming 'liability sinks' for artificial intelligence. *Future Healthcare Journal*, 11(1), 100007. <https://doi.org/10.1016/j.fhj.2024.100007>
- Ledziński, Ł., y Grzešek, G. (2023). Artificial intelligence as an emerging tool for cardiologists. *Medical Sciences Forum*, 21(1), 15. <https://doi.org/10.3390/ECB2023-14339>
- Lin, C.-S., Liu, W.-T., Tsai, D.-J., Lou, Y.-S., Chang, C.-H., Lee, C.-C., Fang, W.-H., Wang, C.-C., Chen, Y.-Y., Lin, W.-S., Cheng, C.-C., Lee, C.-C., Wang, C.-H., Tsai, C.-S., Lin, S.-H., y Lin, C. (2024). AI-enabled electrocardiography alert intervention and all-cause mortality: A pragmatic randomized clinical trial. *Nature Medicine*, 30(5), 1461-1470. <https://doi.org/10.1038/s41591-024-02961-4>
- Lin, T.-H., Chung, H.-Y., Jian, M.-J., Chang, C.-K., Perng, C.-L., Chang, F.-Y., Chen, Y.-H., y Shang, H.-S. (2025). Implementing an AI-enhanced clinical decision support system for *Stenotrophomonas maltophilia*: A survey-based randomized controlled trial of antibiotic precision and impact on survival. *Implementation Science*, 20(1), 47. <https://doi.org/10.1186/s13012-025-01453-4>
- Loftus, T. J., Tighe, P. J., Filiberto, A. C., Efron, P. A., Brakenridge, S. C., Mohr, A. M., Rashidi, P., Upchurch, G. R., Jr., y Bihorac, A. (2020). Artificial intelligence and surgical decision-making. *JAMA Surgery*, 155(2), 148-158. <https://doi.org/10.1001/jamasurg.2019.4917>
- Lorenzini, G., Arbelaes Ossa, L., Shaw, D. M., y Elger, B. S. (2023). Artificial intelligence and the doctor-patient relationship expanding the paradigm of shared decision making. *Bioethics*, 37(5), 424-429. <https://doi.org/10.1111/bioe.13158>
- Mahwish, N., Saherawala, B. A., y Jhancy, M. (2022). Clinical decision making in dysmorphology: Emerging role of artificial intelligence. *British Journal of Healthcare and Medical Research*, 9(5), 366-374. <https://doi.org/10.14738/jbemi.95.13309>
- Mehta, V. (2023). Artificial intelligence in medicine: Revolutionizing healthcare for improved patient outcomes. *Journal of Medical Research and Innovation*, 7(2), e000292. <https://doi.org/10.32892/jmri.292>
- Mingels, S., Piehl, H., Therrien, M., Akhmad, E., van Hienen, A., van Soest, J., Hochstenbach, L., Dekker, A., Damman, O., y Fijten, R. (2026). Understanding clinicians' informational needs for AI-driven clinical decision support systems: Qualitative interview study. *JMIR Medical Education*, 12, e85228. <https://doi.org/10.2196/85228>
- Newman-Toker, D. E., Nassery, N., Schaffer, A. C., Yu-Moe, C. W., Clemens, G. D., Wang, Z., Zhu, Y., Saber Tehrani, A. S., Fanai, M., Hassoon, A., y Siegal, D. (2024). Burden of serious harms from diagnostic error in the USA. *BMJ Quality & Safety*, 33(2), 109-120. <https://doi.org/10.1136/bmjqs-2021-014130>
- Obong'o, C., Njenga, G., Otiangala, D., Jepleting, E., Wairimu, S., Emmanuel-Fabula, M., Kiptinness, S., Korom, R., Taliesin, B., y Mateen, B. A. (2026). Mixed-methods evaluation of clinician experiences and adoption patterns of an EHR-integrated generative AI-based clinical decision support uptake by clinicians in Kenya. *BMJ Digital Health & AI*, 2(1), e000207. <https://doi.org/10.1136/bmjdhai-2025-000207>
- Organización Mundial de la Salud. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. <https://www.who.int/publications/i/item/9789240029200>
- Organización Mundial de la Salud. (2024). *Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models*. <https://www.who.int/publications/i/item/9789240084759>
- Organización Panamericana de la Salud. (2021). *8 principios rectores de la transformación digital del sector de la salud: Un llamado a la acción panamericana* (OPS/EIH/IS/21-0004).
- Organización Panamericana de la Salud. (2025, 18 de noviembre). *Inteligencia artificial y fuerza de trabajo en salud: OPS presenta resultados de un estudio regional que orientará políticas en las Américas*. <https://www.paho.org/es/noticias/18-11-2025-inteligencia-artificial-fuerza-trabajo-salud-ops-presenta-resultados-estudio>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Ramgopal, S., Sanchez-Pinto, L. N., Horvat, C. M., Carroll, M. S., Luo, Y., y Florin, T. A. (2023). Artificial intelligence-based clinical decision support in pediatrics. *Pediatric Research*, 93(2), 334-341. <https://doi.org/10.1038/s41390-022-02226-1>

- Rathore, F. A., y Rathore, M. A. (2023). The emerging role of artificial intelligence in healthcare. *Journal of the Pakistan Medical Association*, 73(7), 1368-1369. <https://doi.org/10.47391/JPMA.23-48>
- Rethlefsen, M. L., Kirtley, S., Waffenschmidt, S., Ayala, A. P., Moher, D., Page, M. J., y Koffel, J. B. (2021). PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Systematic Reviews*, 10(1), 39. <https://doi.org/10.1186/s13643-020-01542-z>
- Santoro, E. (2023). Intelligenza artificiale in medicina: Siamo pronti? [Artificial intelligence in medicine: Are we ready?]. *Recenti Progressi in Medicina*, 114(3), 142-144. <https://doi.org/10.1701/3981.39636>
- Schaffernak, I., Cecil, J., Kleine, A.-K., y Lermer, E. (2025). Sociotechnical influences on the adoption and use of AI-enabled clinical decision support systems in ophthalmology: A theory-based interview study. *BMC Health Services Research*, 25, 1398. <https://doi.org/10.1186/s12913-025-13620-w>
- Shulha, M., Hovdebo, J., D'Souza, V., Thibault, F., y Harmouche, R. (2024). Integrating explainable machine learning in clinical decision support systems: Study involving a modified design thinking approach. *JMIR Formative Research*, 8, e50475. <https://doi.org/10.2196/50475>
- Siden, R., Kerman, H., Gallo, R. J., Cool, J. A., Hom, J., Goh, E., Ahuja, N., Heidenreich, P., Shieh, L., Yang, D., Chen, J. H., Rodman, A., y Holdsworth, L. M. (2025). A typology of physician input approaches to using AI chatbots for clinical decision-making. *npj Digital Medicine*, 9, 14. <https://doi.org/10.1038/s41746-025-02184-y>
- Sorokin, D. Yu., Trufanova, E. S., Rebrova, O. Yu., Bezlepina, O. B., y Laptev, D. N. (2024). Clinical decision support system based on artificial intelligence for adjusting insulin pump parameters in children with type 1 diabetes mellitus. *Diabetes Mellitus (Сахарный диабет)*, 27(3), 242-253. <https://doi.org/10.14341/DM13167>
- Takita, H., Kabata, D., Walston, S. L., Tatekawa, H., Saito, K., Tsujimoto, Y., Miki, Y., y Ueda, D. (2025). A systematic review and meta-analysis of diagnostic performance comparison between generative AI and physicians. *npj Digital Medicine*, 8, 175. <https://doi.org/10.1038/s41746-025-01543-z>
- Tun, H. M., Rahman, H. A., Naing, L., y Malik, O. A. (2025). Trust in artificial intelligence-based clinical decision support systems among health care workers: Systematic review. *Journal of Medical Internet Research*, 27, e69678. <https://doi.org/10.2196/69678>
- Uihlein, A., Beissel, L., Ajlani, A. H., Orzechowski, M., Leinert, C., Kocar, T. D., Pankratz, C., Schuetze, K., Gebhard, F., Steger, F., Fotteler, M. L., y Denking, M. (2024). Expectations and requirements of surgical staff for an AI-supported clinical decision support system for older patients: Qualitative study. *JMIR Aging*, 7, e57899. <https://doi.org/10.2196/57899>
- Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial. *Diario Oficial de la Unión Europea*, L, 2024/1689.
- Vasey, B., Nagendran, M., Campbell, B., Clifton, D. A., Collins, G. S., Denaxas, S., Denniston, A. K., Faes, L., Geerts, B., Ibrahim, M., Liu, X., Mateen, B. A., Mathur, P., McCradden, M. D., Morgan, L., Ordish, J., Rogers, C., Saria, S., Ting, D. S. W., ... McCulloch, P. (2022). Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nature Medicine*, 28(5), 924-933. <https://doi.org/10.1038/s41591-022-01772-9>
- Wang, L., Chen, X., Zhang, L., Li, L., Huang, Y., Sun, Y., y Yuan, X. (2023). Artificial intelligence in clinical decision support systems for oncology. *International Journal of Medical Sciences*, 20(1), 79-86. <https://doi.org/10.7150/ijms.77205>
- Wang, X., y Wang, P. (2025). Clinical decision-making for uveal melanoma radiotherapy: Comparative performance of experienced radiation oncologists and leading generative AI models. *Frontiers in Oncology*, 15, 1605916. <https://doi.org/10.3389/fonc.2025.1605916>
- Wehkamp, K., Krawczak, M., y Schreiber, S. (2023). The quality and utility of artificial intelligence in patient care. *Deutsches Ärzteblatt International*, 120(27-28), 463-469. <https://doi.org/10.3238/arztebl.m2023.0124>
- Whitney, C., Preis, H., y Ramos Vargas, A. (2025). Anticipatory moral distress in machine learning-based clinical decision support tool development: A qualitative analysis. *SSM - Qualitative Research in Health*, 7, 100540. <https://doi.org/10.1016/j.ssmqr.2025.100540>
- Wilhelm, C., Steckelberg, A., y Rebitschek, F. G. (2025). Benefits and harms associated with the use of AI-related algorithmic decision-making systems by healthcare professionals: A systematic review. *The Lancet Regional Health: Europe*, 48, 101145. <https://doi.org/10.1016/j.lanepe.2024.101145>
- Wong, A., Otles, E., Donnelly, J. P., Krumm, A., McCullough, J., DeTroyer-Cooley, O., Pestru, J., Phillips, M., Konye, J., Penzoza, C., Ghous, M., y Singh, K. (2021). External validation of a widely implemented

- proprietary sepsis prediction model in hospitalized patients. *JAMA Internal Medicine*, 181(8), 1065-1070. <https://doi.org/10.1001/jamainternmed.2021.2626>
- Wu, M., Du, X., Gu, R., y Wei, J. (2021). Artificial intelligence for clinical decision support in sepsis. *Frontiers in Medicine*, 8, 665464. <https://doi.org/10.3389/fmed.2021.665464>
- Xu, W., Wang, X., Yang, L., Meng, M., Sun, C., Li, W., Li, J., Zheng, L., Tang, T., Jia, W., y Chen, X. (2024). Consistency of CSCO AI with multidisciplinary clinical decision-making teams in breast cancer: A retrospective study. *Breast Cancer: Targets and Therapy*, 16, 413-422. <https://doi.org/10.2147/BCTT.S419433>
- Yang, Q., Hao, Y., Quan, K., Yang, S., Zhao, Y., Kuleshov, V., y Wang, F. (2023). Harnessing biomedical literature to calibrate clinicians' trust in AI decision support systems. En *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM. <https://doi.org/10.1145/3544548.3581393>
- Yu, Z., Hu, N., Zhao, Q., Hu, X., Jia, C., Zhang, C., Liu, B., y Li, Y. (2025). The willingness of doctors to adopt artificial intelligence-driven clinical decision support systems at different hospitals in China: Fuzzy set qualitative comparative analysis of survey data. *Journal of Medical Internet Research*, 27, e62768. <https://doi.org/10.2196/62768>
- Zhao, Q., Hu, X., Hu, N., Xu, C., y Yu, Z. (2025). Analysis of clinical doctors' adoption willingness pathway for the artificial intelligence-based clinical decision support system (AI-CDSS): A survey study from China. *Digital Health*, 11, 20552076251396984. <https://doi.org/10.1177/20552076251396984>
- Zmudzki, F., Smeets, R. J. E. M., Groenewegen, J. S., y van der Graaff, E. (2025). Machine learning clinical decision support for interdisciplinary multimodal chronic musculoskeletal pain treatment: Prospective pilot study of patient assessment and prognostic profile validation. *JMIR Rehabilitation and Assistive Technologies*, 12, e65890. <https://doi.org/10.2196/65890>