

Aplicación docente de la inteligencia artificial en la educación médica: revisión narrativa del estado actual

Teaching application of artificial intelligence in medical education: a narrative review of the current state

Francisco José Somarriba-López^{1*}  <https://orcid.org/0009-0007-3784-6199>

¹Doctor en Medicina y Cirugía. Docente a tiempo completo de la carrera de Medicina. Facultad de Ciencias Médicas. Universidad Central de Nicaragua. Managua, Nicaragua.

*Autor para la correspondencia. Correo electrónico:  francisco.somarriba@ucn.edu.ni

RESUMEN

Introducción: la inteligencia artificial es una de las tecnologías emergentes más influyentes en los ámbitos de la salud y la educación médica. Posibilita analizar grandes volúmenes de datos educativos o médicos, reconocer tendencias en el desempeño de los estudiantes, y ofrecer experiencias de aprendizaje personalizadas, de acuerdo a las necesidades de cada alumno.

Objetivo: analizar el estado actual de la investigación sobre el uso de la inteligencia artificial como herramienta de apoyo en la educación médica, sus principales aplicaciones, beneficios, desafíos, y perspectivas de integración en entornos universitarios.

Métodos: se realizó una revisión narrativa de la literatura. Para la búsqueda de información se utilizaron las bases de datos *PubMed*, *Scopus*, *Web of Science*, y el buscador *Google Scholar*. Se incluyeron artículos en idiomas español e inglés, publicados entre 2014 y 2025.

Desarrollo: los usos de la inteligencia artificial en la enseñanza de las ciencias básicas médicas se encuentran en sus fases iniciales; no obstante existen ejemplos prometedores. La simulación clínica es uno de los campos más abiertos a la aplicación de esta tecnología en la educación médica, junto a los sistemas automatizados de evaluación y los tutores inteligentes con retroalimentación adaptativa.

Conclusiones: las instituciones deben asumir la integración curricular de estas tecnologías de forma ética, responsable y sostenible; e impulsar un cambio cultural que reconozca sus beneficios, sin perder de vista sus limitaciones y riesgos. Esta revisión aportó una mirada amplia y novedosa sobre la inteligencia artificial en la educación médica, sus fortalezas, y aquellos aspectos que todavía requieren atención.

Palabras clave: algoritmos adaptativos; educación médica; inteligencia artificial; inteligencia artificial generativa; literatura de revisión como asunto; retroalimentación formativa; sistemas inteligentes; tutoría.

ABSTRACT

Introduction: artificial intelligence is one of the most influential emerging technologies in the fields of healthcare and medical education. It enables the analysis of large volumes of educational or medical data, the identification of trends in students' performance, and the delivery of personalized learning experiences tailored to each learner's needs.

Objective: to analyze the current state of research on the use of artificial intelligence as a support tool in medical education, its main applications, benefits, challenges, and prospects for integration in university settings.

Methods: a narrative review of the literature was conducted. The databases *PubMed*, *Scopus*, and *Web of Science*, as well as the *Google Scholar* search engine, were used to retrieve information. Articles published in Spanish and English between 2014 and 2025 were included.

Development: the use of artificial intelligence in teaching basic medical sciences is still in its early stages; however, there are promising examples. Clinical simulation is one of the fields most open to the application of this technology in medical education, along with automated assessment systems and intelligent tutors with adaptive feedback.

Conclusions: institutions should undertake the curricular integration of these technologies in an ethical, responsible, and sustainable manner, and promote a cultural shift that recognizes their benefits without losing sight of their limitations and risks. This review provides a broad and novel perspective on artificial intelligence in medical education, highlighting its strengths as well as the aspects that still require further attention.

Keywords: adaptative algorithms; artificial intelligence; formative feedback; generative artificial intelligence; intelligent systems; medical education; mentoring; review literature as tropic.

Recibido: 04/07/2025

Aprobado: 05/12/2025

Publicado: 16/12/2025

INTRODUCCIÓN

En los últimos años, la inteligencia artificial (IA) se ha posicionado como una de las tecnologías emergentes más influyentes en los ámbitos de la salud y la educación médica. En este contexto, la IA abarca herramientas y técnicas muy diversas: desde algoritmos de aprendizaje automático y modelos de aprendizaje profundo, hasta sistemas de procesamiento de lenguaje natural. Las cuales se utilizan para optimizar la enseñanza y el aprendizaje.⁽¹⁾ Con estas tecnologías es posible analizar grandes volúmenes de datos educativos o médicos, reconocer tendencias en el desempeño de los estudiantes, y ofrecer experiencias de aprendizaje personalizadas, de acuerdo a las necesidades de cada alumno.

El empleo de la IA en la educación médica aporta potencial para transformar las metodologías educativas tradicionales. Posibilita ampliar el acceso a recursos educativos (mediante simulaciones virtuales de pacientes, disponibles en cualquier momento), mejorar la enseñanza de habilidades clínicas y básicas (con la guía de tutores virtuales interactivos), y automatizar procesos evaluativos rutinarios, (de modo que los docentes se centren en aspectos más formativos).⁽²⁾

Las aplicaciones actuales de la IA en la educación médica redundan en mejoras de la enseñanza procedimental, personalización de las experiencias de los estudiantes, y ampliación del alcance de la instrucción, especialmente en especialidades como cirugía y radiología.⁽³⁾ La introducción de modelos de lenguaje de última generación (*ChatGPT*) abre aún más la discusión sobre el papel de la IA en la formación médica; ello, por su capacidad de responder preguntas, generar contenidos educativos, y simular conversaciones de carácter médico convincentes.⁽³⁾

No obstante, junto con las oportunidades, emergen interrogantes y desafíos importantes. La evidencia sobre la eficacia educativa de muchas herramientas de IA, aún es limitada o preliminar, y a menudo es fruto de estudios pilotos a pequeña escala. Al mismo tiempo, existen preocupaciones en torno a la ética, privacidad, equidad, y seguridad al incorporar la IA en entornos formativos. Los educadores médicos reconocen la necesidad de comprender las capacidades y limitaciones de estas herramientas antes de adoptarlas de forma extensiva.⁽²⁾

Se advierte que, sin la preparación adecuada, el uso indiscriminado de la IA podría amenazar la integridad académica o introducir sesgos inadvertidos en los procesos educativos.⁽⁴⁾ Por tanto, es esencial analizar críticamente cómo se utiliza en la educación médica, y cuáles son sus resultados, para proyectar su integración futura de manera responsable.

Por tanto, el objetivo de este trabajo es analizar el estado actual de la investigación sobre el uso de la inteligencia artificial como herramienta de apoyo en la educación médica, sus principales aplicaciones, beneficios, desafíos, y perspectivas de integración en entornos universitarios.

MÉTODOS

Se realizó una revisión narrativa de la literatura. Para la búsqueda de información se utilizaron las bases de datos *PubMed*, *Scopus*, *Web of Science*, y el buscador *Google Scholar*. Se incluyeron artículos en idiomas español e inglés, publicados entre 2014 y 2025, de modo que abarcaran aproximadamente la última década de avances relevantes en la materia.

Se emplearon palabras clave en ambos idiomas. Se combinaron términos provenientes de los Descriptores en Ciencias de la Salud/*Medical Subject Headings* (DeCS/MeSH), y texto libre: *Artificial Intelligence in Medical Education*, *AI in Medical Simulation*, *Intelligent Tutoring Systems in Medical Education*, e Inteligencia artificial en educación médica. Se utilizaron los operadores booleanos Y/O. para relacionar los términos de búsqueda. Y se revisaron manualmente las listas de referencias de artículos relevantes (técnica de búsqueda retrospectiva), para incluir otros estudios acordes a los criterios de inclusión de la investigación.

Los criterios de inclusión fueron: artículos originales, revisiones narrativas o sistemáticas, y estudios experimentales relacionados con la educación médica, publicados entre 2014 y 2025 en revistas

indexadas con revisión por pares, centrados en la aplicación de herramientas de IA en procesos docentes de la medicina, enseñanza de las ciencias básicas, simulación de prácticas médicas, métodos de evaluación o estrategias didácticas innovadoras.

Se excluyeron las publicaciones sin acceso al texto completo, no sometidas a revisión por pares, los estudios que no se centraron en la enseñanza en medicina (IA aplicada sólo a la práctica médica, sin evaluación educativa), y aquellos artículos en idiomas distintos del español e inglés.

Se encontraron 93 referencias potenciales. Tras eliminar publicaciones duplicadas y cribar las restantes en tres etapas (lectura de títulos, revisión de resúmenes, y lectura del texto completo), se seleccionaron 25 artículos. La revisión se enfocó en fuentes documentales caracterizadas por el rigor académico y su relevancia para la educación médica. Se verificó la calidad metodológica de cada una (sin puntuación formal, por el carácter narrativo de la revisión), y se extrajeron los hallazgos clave.

Los estudios seleccionados para la investigación, se organizaron y categorizaron según sus áreas de empleo principales:⁽¹⁾ aplicaciones de IA en la enseñanza de ciencias básicas médicas,⁽²⁾ uso de simuladores inteligentes y aprendizaje adaptativo,⁽³⁾ evaluación automatizada y sistemas de retroalimentación, tutorías inteligentes y asistencia docente.⁽⁴⁾ El análisis crítico realizado posibilitó comparar e integrar los resultados, e identificar beneficios, limitaciones, vacíos de conocimientos, y perspectivas futuras en cada temática.

DESARROLLO

La inteligencia artificial en las ciencias básicas médicas

Los usos de la IA en la enseñanza de las ciencias básicas médicas se encuentran en sus fases iniciales; no obstante existen ejemplos prometedores. En la asignatura Anatomía, las herramientas basadas en IA se utilizan para analizar digitalmente estructuras anatómicas, y crear recursos docentes enriquecidos. En este punto, se deben tener en cuenta las alertas relativas a la relación entre los beneficios y los peligros de integrar IA en la asignatura. Aunque la IA posibilita valerse de tutores virtuales de anatomía y técnicas de identificación automatizada en imágenes, aún no capta plenamente la complejidad espacial y contextual de la anatomía humana.⁽⁵⁾

Los autores del estudio citado,⁽⁵⁾ describieron cinco ejes de discusión alrededor de la incertidumbre de utilizar la IA en la asignatura Anatomía. Si se superan ciertas limitaciones (como la necesidad de curar rigurosamente los datos anatómicos usados para entrenar los algoritmos), la IA podría complementar la enseñanza tradicional mediante imágenes y evaluaciones automatizadas precisas.

En las áreas de patología y ciencias morfológicas, la IA también ha tenido aplicaciones educativas. Un estudio pionero⁽⁶⁾ analizó el comportamiento de los estudiantes ante imágenes histopatológicas; mediante técnicas de visión por computadora, se rastrearon sus patrones de observación de láminas durante exámenes prácticos, y se predijo con un grado alto de exactitud su desempeño en esas evaluaciones. Esto sugiere que la IA puede emplearse para detectar dificultades de aprendizaje en temas básicos (identificación de tipologías de imágenes, o conceptos más problemáticos para un grupo específico de estudiantes); lo cual facilitaría a los docentes ajustar sus estrategias didácticas.

De forma similar, en la asignatura Fisiología médica, existen antecedentes de tutores inteligentes (los sistemas expertos), diseñados para ayudar a los alumnos a razonar sobre los mecanismos homeostáticos. Se documentó⁽⁷⁾ que un tutor automatizado replicó con acierto mayoritario (62-83 % de los casos) las decisiones pedagógicas de tutores humanos, y sostuvo con los estudiantes diálogos didácticos sobre fisiología cardiovascular comparables a los de un instructor humano.

Si bien varios de estos sistemas datan de décadas pasadas, concepciones recientes incorporan el aprendizaje automático para las explicaciones o pistas que se dan a los estudiantes según sus respuestas. Esto demuestra la factibilidad de emular parcialmente la guía de un profesor en entornos básicos, al menos para cierto tipo de contenidos y preguntas estructuradas.

Las respuestas de *ChatGPT* a preguntas de ciencias básicas suelen ser coherentes, pero no infalibles; presentan limitaciones graves especialmente en la interpretación de imágenes (si se trata de identificar estructuras en una imagen histológica) o tablas.⁽³⁾ Se reconoce la IA como un apoyo a la comprensión teórica (respuestas a dudas rápidas, o explicaciones de conceptos fundamentales), pero no puede reemplazar la instrucción formal y su uso requiere supervisión. Es una herramienta valiosa de apoyo al aprendizaje de las ciencias básicas médicas, por su variedad de recursos, pero su incorporación plena en clases teóricas y prácticas de Anatomía y Fisiología implica retos técnicos y metodológicos por superar.

Simuladores inteligentes y aprendizaje adaptativo

La simulación clínica es uno de los campos más abiertos a la aplicación de la IA en la educación médica. Tradicionalmente, la simulación (mediante pacientes estandarizados, maniqués avanzados, o entornos virtuales), reproduce –de manera segura– experiencias de la práctica médica para entrenar a estudiantes y residentes. Con el uso de la IA las simulaciones son más realistas, interactivas, y adaptativas, y la evaluación del desempeño en tiempo real es más objetiva.

En el ámbito de las habilidades y procedimientos quirúrgicos, varias iniciativas han integrado IA con simuladores de realidad virtual. Un ejemplo notable es el *Virtual Operative Assistant (VOA)*, un asistente virtual inteligente diseñado para entrenar en procedimientos quirúrgicos. En un ensayo clínico aleatorizado⁽⁸⁾ con estudiantes de Medicina, aquellos que practicaban con retroalimentación del VOA, tenían mejores desempeños y transferencias de habilidades, comparados con los que eran instruidos de forma remota por un experto humano. La IA les proporcionó comentarios y correcciones durante la simulación quirúrgica; los cuales les posibilitaron ejecuciones superiores de las técnicas operatorias simuladas durante las evaluaciones posteriores.

De forma complementaria, los autores de otro estudio⁽⁹⁾ constataron que el VOA identificó y clasificó correctamente las habilidades de los cirujanos novatos (con diferenciación de los experimentados), según sus métricas de desempeños en una tarea de simulación. Este resultado demostró el potencial de integrar la IA con la realidad virtual, para validar los niveles de competencias técnicas en la enseñanza y práctica quirúrgicas.

La evaluación de destrezas se ha refinado con el uso de la IA. Durante una intervención quirúrgica espinal ejecutada en un entorno de realidad virtual, los algoritmos de aprendizaje automático distinguieron con certeza los diversos niveles de entrenamiento de los practicantes, y correlacionaron sus movimientos y tiempos con sus pericias quirúrgicas.⁽¹⁰⁾

Una aplicación interesante de la IA es el uso de simuladores de pacientes virtuales para entrenar habilidades prácticas como la comunicación entre el médico y el paciente, y la adopción correcta de decisiones médicas. Recientemente, con la aparición de modelos generativos de lenguaje, se han desarrollado pacientes virtuales conversacionales.

Yamamoto y cols.⁽¹¹⁾ utilizaron pacientes simulados con *ChatGPT-4* para que estudiantes de Medicina de cuarto año practicaran entrevistas clínicas. Los resultados fueron contundentes: el grupo que practicó durante un mes con los pacientes virtuales, obtuvo puntajes significativamente más altos; los resúmenes

clínicos elaborados y estructurados a partir de las entrevistas virtuales, fueron más objetivos en comparación con los del grupo de control que usó las tradicionales (28,1/27,1 sobre 31; $p=0,01$). Por otra parte, los alumnos valoraron positivamente la realidad y emotividad de las interacciones con el paciente simulado (mostraba rasgos de personalidad diversos, y estados emocionales según las enfermedades padecidas). De aquí se deduce que los modelos de la IA avanzada pueden complementar o ampliar las prácticas con pacientes estandarizados humanos, posibilitar un entrenamiento flexible (practicar en cualquier momento mediante una plataforma en línea), y retroalimentar al practicante, de forma inmediata, con los resultados de su desempeño en la entrevista.

De manera similar, Maicher y cols.⁽¹²⁾ desarrollaron un sistema de pacientes virtuales con IA para evaluar las habilidades de los estudiantes en la práctica de la anamnesis. El sistema entendía las preguntas de los estudiantes, respondía como lo harían los pacientes, y proveía retroalimentación inmediata sobre puntos clave omitidos o aspectos de comunicación a mejorar. Estas herramientas de simulación conversacional apoyadas por la IA son eficaces para reforzar competencias profesionales como la comunicación, el razonamiento diagnóstico, y la empatía, en un entorno controlado.

La interpretación de exámenes diagnósticos es otra área donde la simulación con IA tiene resultados positivos. Cheng y cols.⁽¹³⁾ incorporaron un módulo de IA en la formación radiológica de estudiantes, concretamente un sistema de aprendizaje automatizado para identificar fracturas de cadera en radiografías de pelvis. Los alumnos que utilizaron esta herramienta mejoraron de forma significativa sus habilidades para reconocer las fracturas, en comparación con aquellos que no emplearon la IA. Esta fue una especie de entrenador durante las simulaciones diagnósticas, resaltó los hallazgos radiológicos, y guió a los estudiantes hacia las características importantes a observar en las imágenes. Esto aceleró su adquisición de habilidades diagnósticas.

Una aplicación de la IA muy interesante es la creación de casos clínicos virtuales inteligentes. En China, la plataforma *Alteach* procesa el lenguaje natural y los datos de historias clínicas reales para generar escenarios médicos interactivos para estudiantes. La práctica con casos virtuales dinámicos mejoró sus habilidades de razonamiento clínico.⁽¹⁴⁾ La plataforma se adaptó a las complejidades de los casos, dio pistas según los progresos de los estudiantes, y simuló las curvas de aprendizaje que tendrían con pacientes reales, pero de forma acelerada.

En general, la evidencia sugiere que la IA puede enriquecer la simulación clínica en la educación médica.⁽¹⁾ Proporciona retroalimentación automatizada y objetiva durante las prácticas (de inmediato señala errores técnicos o de comunicación),⁽²⁾ posibilita prácticas ilimitadas en entornos virtuales seguros con variedad amplia de casos (incluidos escenarios raros que los estudiantes quizás no encontrarían en sus rotaciones prácticas habituales), y facilita monitorear detalladamente sus progresos mediante métricas cuantitativas. Todo ello conforma una experiencia más personalizada.⁽³⁾

Los estudios revisados demuestran las mejorías de desempeño práctico tras intervenciones con simuladores basados en IA, y los grados de satisfacción altos de los usuarios. No obstante, la simulación con IA no es un sucedáneo de la instrucción o supervisión humanas; sí un complemento. Los resultados de las entrevistas clínicas simuladas con *chatbot*, demostraron que los tutores humanos son insustituibles para detectar y emitir matices de empatía, y para validar las conclusiones de los estudiantes.⁽¹⁴⁾ Por otra parte, se reconoce que en las interacciones con pacientes virtuales adicionales, los alumnos ganan confianza y fluidez en sus desempeños sin las presiones inherentes a las evaluaciones formales.⁽¹¹⁾

La IA en la evaluación y sistemas automatizados de calificación

La evaluación del aprendizaje es un componente crítico en la educación médica. Se efectúa mediante exámenes clínicos teóricos y prácticos (objetivos y estructurados) de carácter formativo y acumulativo. La IA ha facilitado la automatización de estas evaluaciones, tanto como la creación de contenidos evaluativos; lo cual hace estos procesos más eficientes, objetivos, y personalizados.

Una línea de investigación interesante y abierta en la actualidad es la generación de preguntas y casos clínicos para exámenes mediante IA. Los modelos de lenguaje natural avanzados posibilitan usar *ChatGPT* para crear ítems de evaluación. Durante un estudio⁽¹⁵⁾ en el *Chaim Sheba Medical Center* (Israel), con *ChatGPT-4* se generaron 210 preguntas de opciones múltiples para exámenes. Médicos especialistas las revisaron a ciegas y constataron que sólo 0,5% contenían errores factuales, y 15,00% requirieron revisiones o ajustes menores. Las preguntas aceptadas abarcaron los contenidos proyectados, y sus niveles de dificultad fueron comparables a los de las redactadas por profesores.

Una comparación de preguntas para exámenes generadas por *ChatGPT* con las elaboradas por profesores de una universidad,⁽¹⁶⁾ evaluadas por un panel independiente, mostró que las diferencias de calidad entre unas y otras no eran significativas. No obstante, la relevancia médica de las confeccionadas por los humanos fue superior.

Estos datos sugieren que la IA podría emplearse para asistir a docentes en la elaboración de bancos de preguntas. Ello ahorraría tiempo en la fase inicial de su redacción, y proporcionaría variantes perfectibles por los propios profesores. Cuando se necesita disponer de un número de casos para prácticas (diagnósticos imagenológicos, intervenciones quirúrgicas simuladas), la IA puede generar escenarios con combinaciones múltiples y diversas de hallazgos patológicos.

Los algoritmos de IA también se utilizan en evaluaciones prácticas. La comprobación de la validez de un sistema automatizado basado en la IA frente a médicos evaluadores entrenados durante una prueba de habilidades clínicas,⁽¹⁷⁾ aportó evidencias de concordancias entre las calificaciones emitidas por la IA y las de los evaluadores humanos. En varios aspectos, la validez de las puntuaciones emitidas por la IA se equiparó a la de las otorgadas por los médicos calificadores.

Este sistema utiliza cámaras o sensores para captar la actuación del estudiante durante la exploración física u otro procedimiento; sus algoritmos identifican las acciones realizadas, y las comparan con las esperadas según las listas de cotejo. Provee retroalimentación, inmediata y detallada, una vez culminada la estación del examen clínico objetivo estructurado (*OSCE*, por sus siglas en inglés), y señala con exactitud los pasos en que los estudiantes cometieron errores u omisiones. Esto último en ocasiones no es posible, por limitaciones, de tiempo en la retroalimentación humana.

De forma análoga, en simuladores laparoscópicos de entrenamiento se incorporaron recursos de IA que evalúan las destrezas técnicas.⁽¹⁸⁾ El sistema entrena a un algoritmo con videos de desempeño experto; posteriormente calcula los niveles de competencias de los aprendices durante la ejecución de exámenes laparoscópicos novedosos y de precisión, y les asigna puntajes de competencias de manera autónoma. Estos desarrollos demuestran que la IA puede aportar objetividad y eficiencia a las evaluaciones prácticas, y reducir las cargas de profesores y tutores al encargarse de la observación minuciosa y puntuación; especialmente en entornos donde se necesitaban dos o más evaluadores para garantizar la confiabilidad.

En la evaluación formativa, la retroalimentación automatizada proporcionada por la IA es completa, porque califica y genera informes personalizados. En Estados Unidos de Norteamérica, la Universidad de Vanderbilt implementó una herramienta de procesamiento de lenguaje natural.⁽¹⁹⁾ La cual analizó y detectó patrones y referencias de competencias geriátricas en las notas clínicas escritas por estudiantes de Medicina durante sus rotaciones. A partir de estos datos, proporcionó a cada estudiante

recomendaciones de aprendizaje específicas. Este sistema identificó las competencias deficientemente desarrolladas por los estudiantes (detectó pocas notas sobre la atención y cuidados a pacientes ancianos con múltiples comorbilidades), y emitió sugerencias o indicó recursos formativos para superar esos vacíos de conocimientos y habilidades.

Booth y cols.,⁽²⁰⁾ utilizaron la IA para examinar las narrativas escritas por médicos evaluadores sobre estudiantes (comentarios de evaluaciones clínicas), con el fin de extraer de forma consistente las subcompetencias o categorías de desempeño. De este modo detectaron con rapidez qué competencias se habían evaluado a cada alumno, y sus niveles. Esto facilitó su seguimiento personalizado mediante currículos basados en las competencias.

No obstante, existen controversias en cuanto al uso de los modelos generativos en los exámenes médicos. Una evaluación del desempeño de *ChatGPT* (*GPT-3.5* y *GPT-4*) en exámenes de la carrera de Medicina⁽³⁾ tuvo resultados sorprendentes. Las puntuaciones de *ChatGPT-4* fueron equivalentes o superiores al umbral de aprobación en la mayoría de las pruebas estandarizadas de las especialidades. Concretamente en el examen de la licencia médica estadounidense, más de 80% de las preguntas generadas por el modelo recibieron puntuaciones dentro del rango de aprobación.⁽²¹⁾

Estos estudios no implican una aplicación educativa *per se* (sólo evalúan la IA como si fuese un estudiante), pero sus implicaciones son relevantes. Demuestran que los bancos de preguntas tradicionales pueden ser resueltos por IA con una tasa de efectividad alta; lo cual conduce a repensar los diseños de las evaluaciones para medir capacidades humanas únicas (razonamiento clínico contextual, comunicación con pacientes reales). Por otro lado, sugieren que en un futuro próximo los estudiantes podrían usar las herramientas de IA como apoyos en su preparación para exámenes; con la garantía de acceder a explicaciones o justificaciones de las respuestas.

Se debe tener cuenta que en esos estudios se identificaron limitaciones de la IA para determinados contextos.⁽³⁾ *ChatGPT-4* tuvo dificultades para responder preguntas que involucraban imágenes, tablas, o diagramas clínicos. En algún examen especializado su versión previa (*GPT-3.5*) reprobó preguntas elaboradas en otro idioma, o con formulaciones muy contextuales.⁽²¹⁾

Esto indica que si bien la IA es potente para procesar información textual y hechos médicos, no reemplaza la experiencia clínica ni el razonamiento de base visual que los médicos en formación deben dominar. En el terreno evaluativo la IA es más una aliada que un sustituto; puede generar y revisar preguntas,

corregir ciertas respuestas, y detectar patrones en los desempeños de los estudiantes. Pero también es un factor que obliga a los profesores a innovar en las evaluaciones, aportar aquellos elementos que la IA aún no puede simular (como la interacción humana real).

Tutores inteligentes y retroalimentación adaptativa

Los tutores inteligentes o sistemas de tutoría inteligente (*ITS*, por sus siglas en inglés) son plataformas de enseñanza automatizada, basadas en IA. Instruyen y retroalimentan a los usuarios de forma personalizada y en tiempo real, para lo cual simulan la guía proporcionada por un tutor humano. En estos sistemas, la IA adapta la presentación de contenidos y la retroalimentación a las necesidades individuales de los estudiantes.⁽²²⁾

Los sistemas de tutoría inteligente se emplean en áreas diversas de la educación médica, como las ciencias básicas, y el desarrollo de habilidades clínicas y de razonamiento diagnóstico.⁽²²⁾ Un aspecto clave es la retroalimentación adaptativa, mediante la cual el sistema analiza las respuestas o acciones de los estudiantes, y les proporciona consejos, correcciones, o pistas ajustadas a sus niveles de desempeño. Ello personaliza sus experiencias educativas y los capacita para el aprendizaje autodirigido y el pensamiento crítico.

Los sistemas de tutoría inteligente integran distintos algoritmos de la IA para modelar los conocimientos médicos, los progresos de los alumnos, y las estrategias pedagógicas. Las técnicas más utilizadas para adaptar la instrucción son los sistemas basados en reglas y redes bayesianas, los métodos de minería de datos, y el aprendizaje automático.⁽²³⁾

Los procesamientos de lenguaje natural y de redes neuronales profundas son funciones incorporadas a la IA recientemente. A la par, se exploran las posibilidades de usar tutores conversacionales basados en modelos de lenguajes grandes para responder preguntas, generar explicaciones personalizadas, y simular pacientes virtuales capaces de entablar diálogos dinámicos.⁽³⁾ Con esta combinación de técnicas el tutor inteligente puede construir modelos de los estudiantes (seguir los progresos y deficiencias de sus conocimientos), o modelos expertos (conocimientos médicos adecuados), y decidir la mejor intervención didáctica en cada caso (dar pistas, corregir, o plantear un nuevo reto).

Los tutores inteligentes para colaboración en grupo y resolución de problemas clínicos, y los agentes conversacionales robóticos son dos herramientas de IA desarrolladas recientemente. Un estudio del año en curso,⁽²⁴⁾ en el Instituto Karolinska (Suecia), integró un robot social controlado por IA, como paciente

virtual con el que los estudiantes dialogaron para practicar el razonamiento clínico. Los estudiantes notaron que –comparada con una plataforma convencional de paciente virtual– esta modalidad aumentó la sensación de autenticidad del encuentro simulado entre el médico y el enfermo, y la percepción del aprendizaje. Estas innovaciones (todavía en etapas iniciales de desarrollo), abren las puertas a experiencias formativas inmersivas. Su retroalimentación adaptativa proviene de avatares o *chatbots* interactivos, que enriquecen el aprendizaje con la comunicación natural y unos escenarios más realistas. De la evidencia científica acumulada entre 2014 y 2025 se deduce que los tutores inteligentes con retroalimentación adaptativa pueden mejorar significativamente los resultados del aprendizaje en la carrera de Medicina. Los estudiantes logran mayores rendimientos académicos, reflejados en calificaciones mejores en los exámenes teóricos y prácticos, y en la ejecución de procedimientos técnicos.⁽¹⁴⁾ Se constata que la transferencia y retención de habilidades es elevada, debido a que el entrenamiento con sistemas de tutoría inteligente favorece el desempeño incluso ante casos nuevos o simulaciones realistas.⁽⁸⁾ La retroalimentación inmediata y consistente ayuda a corregir los errores al momento y refinar las estrategias de razonamiento o técnicas antes de que se arraiguen prácticas erradas.^(8,25) La autonomía y personalización en el aprendizaje permiten adaptar la enseñanza al ritmo y necesidades individuales; esto último es valioso especialmente en entornos clínicos donde los estudiantes tienen niveles diversos de conocimientos previos.⁽⁸⁾

Las tecnologías de tutores inteligentes y retroalimentación adaptativa transforman positivamente la educación médica, mediante la instrucción personalizada a escala, y la réplica (en grados distintos) de la experticia del docente humano. Las plataformas basadas en IA implementadas en universidades y centros de simulación de vanguardia, han contribuido a mejorar las competencias teóricas y prácticas de los estudiantes de Medicina en condiciones controladas. No obstante, quedan retos por superar en cuanto a su integración curricular, aceptación por los profesores, y consideraciones éticas y de costos a tener en cuenta.

La incorporación estratégica de los tutores virtuales puede enriquecer la formación médica, optimizar el proceso de enseñanza y aprendizaje, y contribuir potencialmente a egresar profesionales de la salud mejor preparados. Las investigaciones futuras deberán centrarse en evaluar su impacto a largo plazo en el desempeño profesional de los estudiantes, y explorar formas de combinar la tutoría inteligente con la enseñanza tradicional para conservar lo mejor de ambos enfoques.

Debido a su naturaleza, la presente revisión narrativa tiene limitaciones. Al no ser sistemática, presenta sesgos en cuanto a las fuentes elegidas por el autor. Lo cual intentó mitigar con una búsqueda amplia en múltiples bases de datos y la aplicación explícita de criterios de inclusión y exclusión. Se pueden haber omitido investigaciones relevantes, no identificadas en la búsqueda o publicadas en otros idiomas. Los estudios seleccionados se evaluaron cualitativamente de forma general, sin utilizar herramientas estandarizadas de valoración crítica; esto podría influir en la interpretación de los hallazgos. Por otra parte, la IA evoluciona de forma rápida, por lo cual existen brechas inevitables entre la generación de evidencias nuevas y su divulgación en publicaciones académicas; por ello, puede que algunos desarrollos muy recientes no estén aún reportados en la literatura.

Pese a estas limitaciones, se considera que la revisión ofrece un panorama representativo y suficientemente actualizado (más de 70 % de sus referencias son de los últimos cinco años) de las principales tendencias, ventajas, y desafíos en el uso de la IA en la educación médica.

CONCLUSIONES

Se constata un aumento del empleo de herramientas basadas en inteligencia artificial en la formación profesional en ciencias de la salud. Su uso en las ciencias básicas facilita que los estudiantes comprendan fenómenos que antes percibían abstractos. En los espacios de simulación clínica, los escenarios que crea son más realistas y seguros (es posible equivocarse y aprender sin riesgos). En las evaluaciones, agiliza el seguimiento de los progresos individuales. La integración de la inteligencia artificial en los sistemas de apoyo académico, posibilita una orientación continua y ajustada a las necesidades de los estudiantes; lo cual redundará en su rendimiento, participación, y confianza durante el aprendizaje. A la vez, los docentes pueden dedicar más tiempo a tareas que requieren criterios humanos y enfoques pedagógicos. Las instituciones deben asumir la integración curricular de estas tecnologías de forma ética, responsable y sostenible; e impulsar un cambio cultural que reconozca sus beneficios, sin perder de vista sus limitaciones y riesgos. Esta revisión reunió y analizó críticamente la evidencia científica, y aportó una mirada amplia y novedosa sobre la inteligencia artificial en la educación médica, sus fortalezas, y aquellos aspectos que todavía requieren atención.

REFERENCIAS BIBLIOGRÁFICAS

1. Hernández-Rincón EH, Jiménez D, Chavarro-Aguilar LA, Pérez-Flórez JM, Romero-Tapia ÁE, Jaimes-Peñuela CL. Mapping the use of artificial intelligence in medical education: a scoping review. BMC Med Educ [Internet]. 2025 [citado 8 May 2025];25(1):526. Disponible en: <https://link.springer.com/content/pdf/10.1186/s12909-025-07089-8.pdf?>
2. Shaw K, Henning MA, Webster CS. Artificial intelligence in medical education: a scoping review of the evidence for efficacy and future directions. Med Sci Educ [Internet]. 2025 [citado 8 May 2025];35(3):1803-1816. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC12228863/pdf/40670_2025_Article_2373.pdf
3. Hallquist E, Gupta I, Montalbano M, Loukas M. Applications of artificial intelligence in medical education: a systematic review. Cureus [Internet]. 2025 [citado 8 May 2025];17(3):e79878. Disponible en: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11872247/pdf/cureus-0017-00000079878.pdf>
4. Patino GA, Amiel JM, Brown M, Lypson ML, Chan TM. The promise and perils of artificial intelligence in health professions education practice and scholarship. Acad Med [Internet]. 2024 [citado 8 May 2025];99(5):477-81. Disponible en: https://www.researchgate.net/profile/Megan-Brown-31/publication/377662752_The_Promise_and_Perils_of_Artificial_Intelligence_in_Health_Professions_Education_Practice_and_Scholarship/links/65b8e1f634bbff5ba7d9b909/The-Promise-and-Perils-of-Artificial-Intelligence-in-Health-Professions-Education-Practice-and-Scholarship.pdf
5. Lazarus MD, Truong M, Douglas P, Selwyn N. Artificial intelligence and clinical anatomical education: promises and perils. Anat Sci Educ [Internet]. 2022 [citado 8 May 2025];17(2):249-62. Disponible en: <https://anatomypubs.onlinelibrary.wiley.com/doi/epdf/10.1002/ase.2221>
6. Walkowski S, Lundin M, Szymas J, Lundin J. Exploring viewing behavior data from whole slide images to predict correctness of students' answers during practical exams in oral pathology. J Pathol Inform [Internet]. 2015 [citado 8 May 2025];6(1):28. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2153353922004837>
7. Suebnukarn S, Haddawy P. A Bayesian approach to generating tutorial hints in a collaborative medical problem-based learning system. Artif Intel IMed [Internet]. 2006 [citado 8 May 2025];38(1):5-24.

Disponible

en:

<https://www.sciencedirect.com/science/article/abs/pii/S0933365705000898?via%3Dihub>

8. Fazlollahi AM, Bakhaidar M, Alsayegh A, Yilmaz R, Winkler-Schwartz A, Mirchi N, et al. Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students. JAMA Netw Open [Internet]. 2022 [citado 8 May 2025];5(2):e2149008. Disponible en:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC8864513/>

9. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. Plos One [Internet]. 2020 [citado 8 May 2025];15(2):e0229596. Disponible en:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0229596>

10. Bissonnette V, Mirchi N, Ledwos N, Alsidieri G, Winkler-Schwartz A, Del Maestro RF. Artificial intelligencedistinguishessurgical training levels in a virtual realityspinaltask. J Bone Joint Surg Am [Internet]. 2019 [citado 8 May 2025];101(23):e127. Disponible en:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC7406145/pdf/jbjsam-101-e127.pdf>

11. Yamamoto A, Koda M, Ogawa H, Miyoshi T, Maeda Y, Otsuka F, et al. Enhancing medical interview skills through AI simulated patient interactions: non-randomized controlled trial. JMIR Med Educ [Internet]. 2024 [citado 8 May 2025];10:e58753. Disponible en:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC11459107/>

12. Maicher KR, Zimmerman L, Wilcox B, Liston B, Cronau H, Macerollo A, et al. Using virtual standardized patients to accurately assess information gathering skills in medical students. Med Teach [Internet]. 2019 [citado 8 May 2025];41(9):1053-9. Disponible en:

<https://www.tandfonline.com/doi/full/10.1080/0142159X.2019.1616683#>

13. Cheng CT, Chen CC, Fu CY, Chaou CH, Wu YT, Hsu CP, et al. Artificial intelligence-based education assists medical students' interpretation of hip fracture. Insights Imaging [Internet]. 2020 [citado 8 May 2025];11(1):119. Disponible en:

https://pmc.ncbi.nlm.nih.gov/articles/PMC7683624/pdf/13244_2020_Article_932.pdf

14. Wang M, Sun Z, Jia M, Wang Y, Wang H, Zhu X, et al. Intelligent virtual case learning system based on real medical records and natural language processing. BMC Med Inform Decis Mak [Internet]. 2022

[citado 8 May 2025];22(1):60. Disponible en:

https://pmc.ncbi.nlm.nih.gov/articles/PMC8895690/pdf/12911_2022_Article_1797.pdf

15. Klang E, Portugez S, Gross R, Kassif LR, Brenner A, Gilboa M, et al. Advantages and pitfalls in utilizing artificial intelligence for crafting medical examinations: a medical education pilot study with GPT-4. BMC Med Educ [Internet]. 2023 [citado 8 May 2025];23(1):772. Disponible en:

https://pmc.ncbi.nlm.nih.gov/articles/PMC10580534/pdf/12909_2023_Article_4752.pdf

16. Cheung BH, Lau GK, Wong GT, Lee EY, Kulkarni D, Seow CS, et al. ChatGPT versus human in generating medical graduate exam multiple choice questions—a multinational prospective study (Hong Kong S.A.R., Singapore, Ireland, and the United Kingdom). Plos One [Internet]. 2023 [citado 8 May 2025];18(8):e0290691. Disponible en:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC10464959/pdf/pone.0290691.pdf>

17. Johnsson V, Søndergaard MB, Kulasegaram K, Sundberg K, Tiblad E, Herling et al. Validity evidence supporting clinical skills assessment by artificial intelligence compared with trained clinician raters. Med Educ [Internet]. 2023 [citado 8 May 2025];58(1):105-17. Disponible en:

<https://asmepublications.onlinelibrary.wiley.com/doi/epdf/10.1111/medu.15190>

18. Ryder CY, Mott NM, Gross CL, Anidi C, Shigut L, Bidwell SS, et al. Using artificial intelligence to gauge competency on a novel laparoscopic training system. J Surg Educ [Internet]. 2023 [citado 8 May 2025];81(2):267-274. Disponible en:

<https://www.sciencedirect.com/science/article/abs/pii/S1931720423003811?via%3Dihub>

19. Denny JC, Spickard A, Speltz PJ, Porier R, Rosenstiel DE, Powers JS. Using natural language processing to provide personalized learning opportunities from trainee clinical notes. J Biomed Inform [Internet]. 2015 [citado 8 May 2025];56:292-9. Disponible en:

<https://www.sciencedirect.com/science/article/pii/S1532046415001148?via%3Dihub>

20. Booth GJ, Ross B, Cronin WA, McElrath A, Cyr KL, Hodgson JA, et al. Competency-Based assessments: leveraging artificial intelligence to predict subcompetency content. Acad Med [Internet].

2022 [citado 8 May 2025];98(4):497-504. Disponible en:

https://journals.lww.com/academicmedicine/abstract/2023/04000/competency_based_assessments_leveraging.20.aspx

21. Knoedler L, Alfertshofer M, Knoedler S, Hoch CC, Funk PF, Cotofana S, et al. Pure wisdom or potemkin villages? A comparison of chatgpt 3.5 and chatgpt 4 on USMLE step 3 style questions: quantitative analysis. JMIR Med Educ [Internet]. 2024 [citado 8 May 2025];10:e51148. Disponible en: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10799278/>
22. Wang H, Tlili A, Huang R, Cai Z, Li M, Cheng Z, et al. Examining the applications of intelligent tutoring systems in real educational contexts: a systematic literature review from the social experiment perspective. Educ Inf Technol [Internet]. 2023 [citado 8 May 2025];28(7):9113-148. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC9825070/pdf/10639_2022_Article_11555.pdf
23. Mousavinasab E, Zarifsanaiey N, R Niakan-Kalhari SR, Rakhshan M, Keikha L, Ghazi-Saeedi M. Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. Interact Learn Environ [Internet]. 2018 [citado 8 May 2025];29(1):142-63. Disponible en: <https://www.tandfonline.com/doi/epdf/10.1080/10494820.2018.1558257>
24. Borg A, Georg C, Jobs B, Huss V, Waldenlind K, Ruiz M, et al. Virtual patient simulations using social robotics combined with large language models for clinical reasoning training in medical education: mixed methods study. J Med Internet Res [Internet]. 2025 [citado 8 May 2025];27:e63312. Disponible en: <https://www.jmir.org/2025/1/e63312/>
25. Feyzi-Behnagh R, Azevedo R, Legowski E, Reitmeyer K, Tseytlin E, Crowley RS. Metacognitive scaffolds improve self-judgments of accuracy in a medical intelligent tutoring system. InstrSci [Internet]. 2013 [citado 8 May 2025];42(2):159-81. Disponible en: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3923630/pdf/nihms466801.pdf>

Conflicto de intereses

El autor declara que no existen conflictos de intereses.

Contribución del autor

Francisco José Somarriba-López: conceptualización, análisis formal, investigación, metodología, recursos, redacción, revisión, y edición.

Financiación

Universidad Central de Nicaragua, Managua, Nicaragua.