InvestiGO
Revista Científica Multidisciplinaria
ISSN 2953-6367
http://revistainvestigo.com

Octubre 2026
Vol. 7 No. 21, PP. 394-406
https://doi.org/10.56519/mrth4j35

INTELIGENCIA ARTIFICIAL GENERATIVA EN LA FORMACIÓN DE PROFESIONALES DE LA SALUD: TRANSFORMACIÓN DE COMPETENCIAS, EVALUACIÓN DEL APRENDIZAJE Y DESAFÍOS ÉTICOS

GENERATIVE ARTIFICIAL INTELLIGENCE IN THE TRAINING OF HEALTHCARE PROFESSIONALS: TRANSFORMATION OF SKILLS, LEARNING ASSESSMENT, AND ETHICAL CHALLENGES

Gina Alexandra Asimbaya Pachacama[1], Leonardo Lanzee Mendoza Jiménez[2], Ivanova del Cisne Zúñiga Román[3], Joselyn Lissette Martinez Altamirano[4]

{gina.asimbaya@quito.gob.ec[1], lic.lanzee.mendoza@hotmail.com[2], ivacisnezr@hotmail.com[3], jossl_5@hotmail.com[4]}

Fecha de recepción: 02/09/2026 / Fecha de aceptación: 08/10/2026 / Fecha de publicación: 14/10/2026

RESUMEN: La inteligencia artificial generativa (IAG) se ha introducido rápidamente en la formación de los profesionales de la salud, reconvirtiendo los modos de enseñar, evaluar y adquirir las competencias clínicas. Se trata de analizar, mediante una revisión narrativa de la literatura científica indexada publicada entre el 2016 y 2026, la evidencia existente sobre la transformación de las competencias, la evaluación del aprendizaje y los dilemas éticos relacionados con la IAG en la educación del profesional de la salud. Para eso, se llevó a cabo una revisión de la literatura de tipo narrativo, siguiendo las directrices PRISMA para la identificación y selección de las fuentes mediante búsqueda en PubMed/MEDLINE, Scopus, Web of Science y SciELO, así como artículos originales, revisiones sistemáticas y de alcance publicados en revistas indexadas. Se encontraron tres ejes temáticos prevalentes: la transformación de las competencias digitales, clínicas y de pensamiento crítico; el uso de la IAG en la evaluación del aprendizaje, incluida su manera de desempeñarse en los exámenes de licenciamiento médico; y los desafíos éticos que afectan la integridad académica, la parcialidad algorítmica, la protección de datos y la despersonalización del cuidado. Los modelos de lenguaje valorados alcanzaron niveles de desempeño similares a la o superiores a la línea de corte necesaria para pasar los exámenes estandarizados, aunque subsisten limitaciones en relación con la validez, fiabilidad y preparación docente. La IAG supone una herramienta de potencial transformador en la educación de los profesionales de la salud, aunque su integración

[1]Unidad Metropolitana de Salud Centro, Secretaría de Salud, Municipio del Distrito Metropolitano de Quito – Ecuador, https://orcid.org/0009-0001-4510-9822
[2]Investigador independiente – México Cdmx, https://orcid.org/0009-0005-3322-6916
[3]Carrera de Laboratorio Clínico, Facultad de la Salud Humana, Universidad Nacional de Loja, Loja, Ecuador– Ecuador, https://orcid.org/0003-3690-7686
[4]Escuela de posgrado, Doctorado en Ciencias Biomédicas, Instituto Universitario Italiano de Rosario (IUNIR), Rosario S2000CTT, Argentina, https://orcid.org/0000-0002-0625-1930

responsable precisará de marcos éticos explícitos, alfabetización digital del profesorado y del alumnado y maneras de evaluar que conserven el razonamiento clínico humano.

Palabras clave: Inteligencia artificial, educación médica, evaluación educacional, ética, personal de salud, competencia clínica

ABSTRACT: Generative artificial intelligence (GAI) has been rapidly integrated into the training of healthcare professionals, transforming the ways in which clinical competencies are taught, assessed, and acquired. The aim of this study is to analyze, through a narrative review of indexed scientific literature published between 2016 and 2026, the existing evidence on the transformation of competencies, the assessment of learning, and the ethical dilemmas related to GAI in health professional education. To this end, a narrative literature review was conducted, following the PRISMA guidelines for the identification and selection of sources through searches in PubMed/MEDLINE, Scopus, Web of Science, and SciELO, as well as original articles, systematic reviews, and overview articles published in indexed journals. Three prevailing thematic areas were identified: the transformation of digital, clinical, and critical thinking skills; the use of AI in learning assessment, including its performance on medical licensing exams; and the ethical challenges affecting academic integrity, algorithmic bias, data protection, and the depersonalization of care. The language models evaluated achieved performance levels similar to or higher than the cutoff score required to pass standardized exams, although limitations remain regarding validity, reliability, and faculty preparation. The IAG represents a potentially transformative tool in the education of health professionals, although its responsible integration will require explicit ethical frameworks, digital literacy among faculty and students, and assessment methods that preserve human clinical reasoning.

Keywords: Artificial intelligence, medical education, educational assessment, ethics, healthcare personnel, clinical competence

INTRODUCCIÓN

La inteligencia artificial (IA) ha pasado de ser una tecnología emergente para considerarse uno de los componentes estructurales de los sistemas de salud contemporáneo en la última década, con aplicaciones que van desde el diagnóstico por imagen, por ejemplo, hasta la gestión administrativa de las instituciones sanitarias. En el marco de este proceso, la entrada de los LLM (Language Model) y las herramientas de inteligencia artificial generativa (IAG), que son como los LLM pero que son también capaces de generar texto, imágenes, casos clínicos simulados o retroalimentación personalizada a partir de instrucciones en lenguaje natural, ha supuesto un cambio especialmente profundo en el ámbito de la educación de los profesionales de la salud (1), (2) . Instituciones académicas de todo el mundo han empezado a usar estas herramientas en sus procesos de enseñanza, de evaluación o de formación continuada, planteando a la vez expectativas de los beneficios que pueden traer para mejorar la calidad formativa y plantear interrogantes sobre sus límites y riesgos (1).

Las evidencias acumuladas en los años recientes dan cuenta de la importancia de este fenómeno. Así, en una revisión de amplitud que resumía 278 publicaciones sobre inteligencia artificial en educación médica, se documentó que la mayor parte de los trabajos están focalizados en Norteamérica y Europa, y que sus aplicaciones van desde procesos de admisión, enseñanza, evaluación y razonamiento clínico (1). Una revisión sistemática que se focalizaba concretamente en herramientas bajo el concepto de IAG nos indicaba que la generación de contenido educativo (casos clínicos, planes de cuidado, preguntas de evaluación) representa uno de los usos más comunes que se recogen en la literatura, con beneficios documentados en alfabetización digital y autoeficacia del estudiantado (3), (4) . Todo esto parece evidenciar que la IAG no automatiza únicamente tareas administrativas, sino que interviene en la construcción de competencias profesionales. La transformación de competencias supone, precisamente, uno de los ejes más relevados.

Las clásicas estructuras de las competencias médicas basadas en el razonamiento clínico, la comunicación y el profesionalismo vienen siendo cuestionadas por la necesidad de incorporar una nueva forma de alfabetización, la alfabetización en inteligencia artificial, definida como la capacidad para usar, interpretar y supervisar de manera crítica los productos de tales sistemas (2), (5). Varios autores han documentado que los alumnos de medicina, enfermería, farmacia y odontología utilizan cada vez con más frecuencia asistentes conversacionales para resolver casos clínicos, escribir notas de evolución a modo de simulacros, preparar exámenes de certificación, etc., lo que está llevando a rediseñar los planes de estudio en cuanto a la formación híbrida entre el conocimiento disciplinar y la competencia digital (5) .

En lo que respecta a la evaluación del aprendizaje, la evidencia es muy sólida. En estudios iniciales, se observó que ChatGPT podía lograr el umbral de aprobación o aproximarse a este umbral en el United States Medical Licensing Examination (USMLE), lo que puede generar controversia en el diseño de instrumentos de evaluación (6), (7) , si bien estudios posteriores que utilizaban versiones más completas del modelo reportaron niveles de precisión superiores al 85 % para las preguntas de tipo USMLE Step 1 y una homogeneidad de resultados en las diferentes disciplinas8. Resultados similares se han documentado en especialidades como neurocirugía, donde los modelos de lenguaje de última generación han superado a usuarios humanos en la resolución de bancos de preguntas de preparación para juntas orales (8). Estos resultados dan lugar a una pregunta clave a tener en cuenta desde la pedagogía de las ciencias de la salud: si un sistema de IAG es capaz de superar exámenes que están pensados para certificar la competencia clínica, ¿qué es lo que entonces debe ser evaluado en la formación de un profesional de la salud en el siglo XXI? El cuarto eje central son los desafíos éticos asociados a la forma en la que estas herramientas pueden influir en la formación de los/las profesionales de la salud.

Una serie de estudios advierte sobre los riesgos relacionados con la integridad académica, el plagio o, incluso, aquella dependencia excesiva de estas herramientas que puede mermar el desarrollo del pensamiento crítico y del razonamiento clínico autónomo (9), (10). Otras preocupaciones van relacionadas con el sesgo algorítmico; la protección de datos sensibles de pacientes y estudiantes; la IAG puede llegar a reproducir o multiplicar inequidades estructurales que ya se encuentran en los datos existentes (11) . En la radiología, la dermatología y la medicina

nuclear, diferentes autores han indicado que la falta de marcos regulatorios explicativos y de programas docentes estructurados para la formación del profesorado permite que las tecnologías tengan un uso poco responsable en la incorporación de estas en la formación de los planes de estudio (12), (13).

En el ámbito permitido de Latinoamérica, la evidencia es todavía muy germinal pero creciente. Recientes estudios realizados en Ecuador, en México, El Salvador y en otros países desarrollados de la región han detectado una adopción acelerada de la IAG entre los docentes y estudiantes de ciencias de la salud, así como aún con brechas importantes en lo que respecta la alfabetización digital y a la existencia de línea. Las revisiones del alcance centradas en la formación del profesorado entre las ciencias de la educación latinoamericanas son coincidentes en afirmar que la realidad de Latinoamérica es la integración de herramientas sin marcos éticos y pedagógicos propios, en muchos casos, replicando modos de hacer pedagógicos construidos a partir de los contextos anglosajones (14). Esto cobra especial relevancia en países como Ecuador, que se enfrenta además a retos estructurales para el acceso equitativo a la tecnología en la formación de los profesionales de la salud.

A pesar del aumento en la producción científica relacionada con la inteligencia artificial en la educación de las ciencias de la salud continúa una fragmentación conceptual; una gran parte de la literatura profundiza de manera independiente la dimensión técnica de los modelos, la dimensión pedagógica de la transformación del currículo o la dimensión ética de su implementación sin realizar una síntesis integradora que empodere a los formadores de los profesionales de la salud a tomar decisiones informadas. Esta fragmentación es el problema que guió esta investigación: la necesidad de sistematizar la evidencia de una manera crítica y actualizada sobre cómo la IAG transforma las competencias profesionales, cómo van evaluando estos aprendizajes y cuáles son los principales retos éticos que se enfrentan respecto a la integración curricular.

La pertinencia de este análisis recae en que los procesos de acreditación y de actualización curricular de los programas de ciencias de la salud deben sustentarse en la evidencia consolidada y no en percepciones subjetivas sobre la utilidad de estas tecnologías. La sistematización de esta evidencia es, además, oportuna para guiar la construcción de políticas institucionales, programas de desarrollo docente y orientaciones para un uso responsable de la IAG en universidades y facultades de las ciencias de la salud, principalmente en contextos como el latinoamericano en el que la producción de orientaciones propias es aún bastante escasa.

Por las razones expuestas, el objetivo de este análisis de revisión fue el de examinar la producción científica indexada, publicada entre 2016 y 2026, sobre la transformación de las competencias profesionales, la evaluación del aprendizaje y los retos éticos relacionados e introducidos por el uso de la Inteligencia Artificial Generativa en la formación de los profesionales de la salud.

MATERIALES Y MÉTODOS

El trabajo en cuestión consiste en una revisión bibliográfica del tipo narrativo, sustentada metodológicamente en las recomendaciones PRISMA (Preferred Reporting Items for Systematic Reviews and Meta Analyses).

No se realizó metaanálisis estadístico propiamente, dado el carácter cualitativo preponderante y de heterogeneidad en diseño y descriptivo de la mayoría de los estudios primarios disponibles.La unidad de análisis estuvo constituida por artículos científicos publicados en revistas indexadas en bases de datos de alto impacto (Scopus, Web of Science, PubMed/MEDLINE y SciELO) sin restricción de país de publicación, priorizando artículos de estudios originales, de revisiones sistemáticas, de revisiones de alcance y de documentos de posicionamiento institucional (guías AMEE y BEME) relacionados con la inteligencia artificial generativa en la educación de profesionales de la salud del periodo 2016-2026.

La investigación se realizó en un entorno documental a través de la consulta remota de bases de datos bibliográficas académicas y de repositorios institucionales de acceso abierto, sin interacción directa sobre sujetos humanos ni sobre instituciones educativas.Las variables que se consideraron en cada fuente analizada fueron las siguientes: año y país donde se publicó, tipo de diseño metodológico, dominio o competencia profesional descrita en la fuente que se analizó, instrumento o estrategia de evaluación del aprendizaje descrita y resultado científico que resume los desafíos éticos que se identifican.

Todas estas variables se presentaron en matrices de extracción de los temas que sirvieron para la síntesis narrativa de los hallazgos. Dado que la revisión en sí misma respondía a un enfoque cualitativo-descriptivo, el análisis se desarrolló mediante estadística descriptiva simple (frecuencias y porcentajes de aparición temáticas) y análisis de contenido de tipo temático-categorial sin el uso de pruebas inferenciales, lo que es completamente acorde con el diseño bibliográfico narrativo de la revisión.

Tabla 1. Estrategia metodológica de búsqueda y selección de la evidencia.

Componente	Descripción
Bases de datos	PubMed/MEDLINE, Scopus, Web of Science (Core Collection) y SciELO
Palabras clave / operadores	("generative artificial intelligence" OR "ChatGPT" OR "large language models") AND ("health professions education" OR "medical education" OR "nursing education") AND ("competencies" OR "assessment" OR "ethics")
Periodo de búsqueda	Enero de 2016 a septiembre de 2026
Idiomas incluidos	Español, inglés y portugués
Criterios de inclusión	Artículos originales, revisiones sistemáticas o de alcance, y guías institucionales (AMEE/BEME) publicados en revistas indexadas, con texto completo disponible y pertinencia temática directa
Criterios de exclusión	Resúmenes de congreso, cartas al editor sin datos originales, publicaciones sin revisión por pares y estudios centrados exclusivamente en IA diagnóstica sin componente educativo
Registros identificados	≈ 210 registros en las cuatro bases consultadas

Registros tras eliminación de duplicados y cribado de título/resumen	68 documentos preseleccionados
Fuentes finalmente incluidas	26 fuentes indexadas con DOI verificado, tras evaluación de calidad, actualidad y pertinencia
Guía de reporte	Adaptación de las fases de identificación, cribado, elegibilidad e inclusión propuestas por PRISMA para revisiones bibliográficas

El proceso de selección siguió un esquema en cuatro fases: identificar los registros en las bases elegidas; eliminar duplicados, proceder al cribado en función de los títulos y los resúmenes; comprobar la elegibilidad mediante la lectura del texto completo en función de criterios de inclusión y exclusión previamente definidos y, finalmente, la inclusión de las fuentes que cumplían los estándares de calidad editorial, actualización (de 2016 a 2026) y correspondencia directa a los tres ejes temáticos que conforman la investigación. Este procedimiento ha permitido reducir el sesgo de selección y garantizar que la síntesis narrativa correspondiente se elaborara únicamente en función de literatura indexada, verificable y por un alto impacto.

RESULTADOS

Realizada la búsqueda y selección de la evidencia, conforme al enfoque PRISMA adaptado a revisiones bibliográficas narrativas 22, se organizaron los 26 documentos incluidos en tres categorías de resultados coherentes con los objetivos planteados: a) la transformación de las competencias profesionales de; b) la práctica de la IAG en los instrumentos de evaluación del aprendizaje; c) los dilemas éticos que aparecen con su aplicación. Después, se sistematizan los hallazgos en tablas descriptivas, para después continuar el análisis correspondiente.

La tabla 2 recoge los principales dominios de la competencia que, según la literatura revisada, se ven modificados por la inclusión de la IAG en los currículos de ciencias de la salud.

Tabla 2. Dominios de competencia transformados por la inteligencia artificial generativa.

Dominio de competencia	Hallazgo principal reportado	Fuente
Alfabetización en IA y pensamiento crítico digital	La incorporación de la IAG exige nuevas competencias de supervisión crítica de los productos generados por los modelos, más allá del dominio disciplinar tradicional	(2), (5)
Razonamiento clínico y elaboración de casos	Uso extendido de la IAG para generar y resolver casos clínicos simulados, con mejoras reportadas en la autoeficacia y la alfabetización digital del estudiantado	(3), (4)
Competencias en farmacia y odontología	La IAG se integra en la elaboración de planes farmacoterapéuticos y en procesos formativos clínicos, generando debates sobre integridad académica y supervisión docente	(15),(12)
ompetencias radiológicas y de imagen diagnóstica	Persisten programas de capacitación breves y poco estructurados en IA aplicada a la formación en radiología, con escasa integración curricular longitudinal	(16) (12)
Competencias docentes	Necesidad de programas estructurados de desarrollo docente y de marcos de progresión de competencias en IA para el profesorado de ciencias de la salud	(17)

Un análisis de estos hallazgos pone de manifiesto que la transformación de las competencias no implica únicamente la adquisición de una serie de habilidades técnicas que permitan el manejo de las herramientas de IAG, sino que exige, además, proporcionar una mayor reconfiguración del perfil profesional esperado, el cual contempla que la supervisión crítica de los productos generados por la máquina tiene un peso específico y relevante junto al conocimiento disciplinar de base (2), (5) . También se observa que la mayoría de la evidencia disponible proviene de la formación médica de pregrado, mientras que disciplinas como la formación enfermera, farmacia, odontología y radiología cuentan con un volumen de evidencia pequeño en comparación con la de otros registros, aunque cada vez más amplio (15), (12). Un hallazgo transversal de interés es la insistencia por parte de distintos autores que afirman que la transformación curricular ha de cumplir una transformación docente ordenada, dado que la competencia para enseñar a supervisar críticamente la IAG da por hecho que el propio profesorado tiene esta competencia bajo control (17), (18).

La Tabla 3 evidencia el conjunto de los principales hallazgos reportados sobre la eficacia de los modelos de lenguaje sobre instrumentos de evaluación aplicados de manera habitual en la formación de los profesionales de la salud.

Tabla 3. Desempeño de modelos de IA generativa en instrumentos de evaluación del aprendizaje.

Estudio	Instrumento evaluado	Principal hallazgo reportado
Kung et al.	USMLE Step 1, 2 y 3	ChatGPT alcanzó o se aproximó al umbral de aprobación en los tres niveles del examen, sin entrenamiento especializado previo
Gilson et al.	USMLE (bancos AMBOSS y NBME)	Desempeño comparable al de un estudiante de tercer año de medicina, con variabilidad relevante entre reactivos e interpretabilidad limitada
Garabet et al.	1300 preguntas de estilo USMLE Step 1	GPT-4 respondió correctamente el 86 % de las preguntas, superando de forma consistente el umbral estimado de aprobación (60 %) en todas las disciplinas evaluadas
Ali et al.	Banco de preparación para juntas orales de neurocirugía	GPT-4 superó el desempeño de ChatGPT y de Google Bard en la resolución de preguntas especializadas de alta complejidad clínica

Los hallazgos resumidos en la Tabla 3 muestran una tendencia nítidamente determinada: los modelos de lenguaje analizados han alcanzado y, en las versiones más actualizadas, han superado con amplitud el umbral mínimo de aprobación para realizar los exámenes de licenciamiento médico (6), (7) . Este hallazgo, que se repite en preguntas de estilo USMLE Step 1 con un porcentaje de corrección de un 86 %8, y ampliado en especialidades quirúrgicas altamente complejas como la neurocirugía (8), se traduce y tiene consecuencias directas sobre la validez de los instrumentos de evaluación que están esencialmente basados en preguntas de opción múltiple, dado que, por el contrario, podrían estar midiendo un tipo de conocimiento, el que es reproducido por sistemas automatizados, y que no se encuentra necesariamente en el corazón de las competencias clínicas complejas el juicio situacional, la comunicación, el manejo de la incertidumbre que nos imaginamos en los profesionales de la salud.

La Tabla 4 es la que nos permite organizar los principales desafíos éticos documentados con respecto al uso de la IAG en la educación de los profesionales de la salud.

Tabla 4. Desafíos éticos asociados al uso de la inteligencia artificial generativa.

Categoría	Descripción del desafío	Fuente
Integridad académica	Riesgo de plagio, dependencia excesiva de las respuestas generadas y erosión del pensamiento crítico y del razonamiento clínico autónomo del estudiantado	(9) , (10)
Sesgo algorítmico y equidad	Posibilidad de que los modelos reproduzcan o amplifiquen inequidades estructurales presentes en los datos utilizados durante su entrenamiento	(11)
Protección de datos y confidencialidad	Preocupaciones relacionadas con el manejo de información sensible de pacientes y estudiantes al utilizar plataformas de IAG de acceso comercial	(19)
Preparación docente y vacíos regulatorios	Ausencia de marcos normativos institucionales claros y de programas de formación docente específica en disciplinas como radiología y dermatología	(12)
Despersonalización del cuidado	Riesgo de que la mediación tecnológica debilite las competencias comunicacionales y empáticas propias de la práctica clínica centrada en la persona	(20)

El análisis de esta categoría de resultados califica de manera holística que la clase de desafíos éticos identificados no son excluyentes tal que vayan en diferentes direcciones, sino que se encuentran reforzándose mutuamente: la ausencia de marcos regulatorios institucionales facilita el uso poco supervisado de la IAG por parte del alumnado, y esto contribuye a incrementar el riesgo de dependencia en exceso y hacer un flaco favor al mantenimiento de la integridad académica (1), (12). Precisamente, diferentes autores coinciden en señalar que la preocupación de fondo no es ya la tecnología, sino la falta de preparación institucional y docente para que puedan hacerse cargo de la integración de sus prácticas reflexivamente y en forma crítica (19) , y el riesgo de que la eficiencia que la IAG proporciona empiece a ir desplazando progresivamente competencias de comunicación y relacionales que son constitutivas del cuidado en salud (20) .

Un subconjunto de la evidencia analizada corresponde específicamente al marco latinoamericano. Algunos estudios realizados en Ecuador y otros países de la región darán cuenta de una adopción creciente de la IAG entre docentes y alumnado de ciencias de la salud, que presentan notables brechas entre lo que se manifiesta saber, las actitudes favorables hacia la tecnología y el uso realizado en el marco de la evaluación y durante la práctica clínica (21). Revisiones de alcance centradas en la formación docente latinoamericana indican que la región aún carece de marcos éticos y pedagógicos de forma que permitan regular el uso de este tipo de herramientas, y existen indicios de que algunas características o rasgos individuales y profesionales como el nivel de formación de posgrado y la exposición previa a la tecnología predicen de forma significativa la adopción por parte del profesorado o el personal clínico (14), (22) . Considerando que la evidencia académica a nivel latinoamericano es escasa, estos resultados refuerzan la pertinencia de generar evidencia localizada que oriente la elaboración de políticas institucionales que tengan en cuenta las condiciones estructurales de la región.

DISCUSIÓN

Los hallazgos de la presente revisión muestran que la inteligencia artificial generativa se ha convertido en un agente de transformación curricular en la formación de profesionales de la salud y se Articula con el hallazgo de la revisión de alcance BEME Guide No. 84; en esta revisión se documentaron 278 publicaciones sobre la materia y se expuso un espectro de aplicaciones que abarca desde los procesos de admisión, pasando por la evaluación, hasta llegar al razonamiento clínico (1). Esta correspondencia indica que la transformación identificada en la presente síntesis no puede considerarse un hallazgo dispar, sino una tendencia consolidada y transversal a varias disciplinas de la salud, tal como también lo hace la guía AMEE No. 158 sobre el uso ético de la inteligencia artificial en la educación de estas profesiones (2).

Con respecto a la transformación de competencias, los resultados de este estudio son coherentes con lo que se reporta en el ámbito de la enfermería, cuyo uso de chat GPT y herramientas afines permite mejorar indicadores de autoeficacia y alfabetización digital, aunque también anticipa el temor docente a que se presente una sustitución del razonamiento clínico propio por la respuesta elaborada de forma automática (9), (10). Esta tensión entre una potencialidad formativa y una dependencia indeseada es, según los autores de esta revisión, el eje del debate contemporáneo en torno a la IA generativa en la educación en salud y coincide con el también observado en el ámbito de la farmacia, pues la generación asistida de planes terapéuticos plantea interrogantes análogos a los del futuro juicio clínico de las y los profesionales (15), (23).

En lo que se refiere a la evaluación del aprendizaje, los resultados aquí sintetizados, que reflejan una precisión de hasta un 86 % en preguntas de estilo USMLE Step 1, presentan una línea de continuidad con el trayecto de estudios precursores que, desde el año 2023, han puesto de manifiesto una mejora progresiva en el rendimiento de los modelos de lenguaje en exámenes de licencias médicas (6), (24) aunque también hay que matizar esta tendencia pues, como advierten Gilson et al., la variablidad entre reactivos y la escasa interpretabilidad de algunas respuestas generadas por el modelo cuestionan la fiabilidad de este respecto de ser una herramienta de apoyo al aprendizaje autónomo y no sólo alcanzar el aprobado en exámenes estandarizados (7). Esta distinción es fundamental: un rendimiento elevado en pruebas de tipo pregunta opción múltiple no es sinónimo de una correcta capacidad de juicio clínico y de empatía, o de capacidad para la gestión de la incertidumbre, dimensiones estas que los instrumentos de evaluación tradicionales apenas logran recoger en el nivel que pudieran/como deberían y que, en un sentido irónico, la irrupción misma de la IAG pasaría a considerarse de nuevo.

Eticamente, en el ámbito de lo ético, los resultados de este estudio expanden lo que también indicaba Currie, que ya en 2023 avisaba del peligro del deterioro de la honestidad académica como producto del uso de estas herramientas, si no son supervis. La ejecución presente sugiere que ese riesgo no pertenece a una disciplina concreta; más bien, se trata de un fenómeno trasversal también registrado en dermatología, por ejemplo, donde la propia disciplina pone de manifiesto la necesidad de definir límites precisos en el uso de la IAG para la construcción de contenido clínico y educativo (13), así como en la más amplia dimensión de la ética de los modelos

de lenguaje orientados a la salud, donde se hace hincapié en la importancia de mantener la supervisión humana como principio para cualquier implementación responsable (11).

Un hallazgo destacado de esta revisión, poco explorado en las síntesis de alcance general, es la brecha regional detectada en América Latina.Por otro lado, la literatura que puede considerarse anglosajona tiende a centrarse en el rendimiento técnico de los modelos y en el diseño de marcos de competencia; toda la evidencia latinoamericana revisada parte de una premisa completamente distinta, desprendiendo un enfoque de problemas tales como una adopción desigual, una formación docente sin apoyos o la carencia de normativas institucionales propias (21), (14). Tal desacuerdo enfatiza que la agenda de investigación en la región debería versar no ya sobre evaluar el rendimiento de la tecnología, sino sobre establecer capacidades institucionales y marcos normativos adaptados a la disponibilidad de recursos y a las necesidades formativas propias de contextos como el ecuatoriano, tal y como han sugerido estudios recientes sobre los predictores individuales y profesionales de la adopción de la IA en el ámbito médico en la región (22).

Finalmente, uno de los aspectos a mencionar de la literatura encontrada es que converge junto a una recomendación válida común: la necesidad de establecer el desarrollo docente estructurado como una entrada de cara a una integración curricular responsable de la IAG y no una inclusión reactiva o y meramente instrumental en cuanto a la incorporación de la tecnología (17) , (18) Subsistiendo un hecho especialmente importante en instituciones formadoras de profesionales de la salud en contextos de escasos recursos, donde la carencia de programas de formación docente propiciaría el ahondamiento de inequidades formativas existentes en lugar de propiciarlas, tal y como también propone de forma muy clara el análisis crítico sobre la incorporación de la IAG en la educación médica continua de la región (25).

Entre las limitaciones de la presente revisión se encuentran las de presentación narrativa que no permiten ofrecer estimaciones del efecto en su agregado, como las que brindaría un metanálisis, o la velocidad dilatante de la que están dotadas las tecnologías analizadas a partir de consideraciones que pueden volver las poblaciones y los hallazgos de rendimiento técnico dentro de un periodo relativamente breve de tiempo. Sin embargo, la triangulación de fuentes de alto impacto publicadas en un periodo de diez años y procedentes de disciplinas y regiones geográficas diferenciadas ofrece un poderoso apoyo para la síntesis conceptual de la que hemos intentado dar cuenta aquí.

CONCLUSIONES

La evidencia científica indexada en la que se fundamenta este texto escrito entre 2016 y 2026, pone de manifiesto que la inteligencia artificial generativa transforma de manera sustancial las competencias exigidas a la formación del profesional de la salud, al promover la incorporación de la alfabetización crítica en IA como parte del bagaje del profesional de la salud junto con el conocimiento disciplinar tradicional y al transformar la manera en que se configura y se ejercita el razonamiento clínico y las competencias comunicacionales.

En lo que respecta a la evaluación del aprendizaje, los modelos de lenguaje generativo han alcanzado niveles de desempeño iguales o superiores al umbral de aprobación en una batería de exámenes de licenciamiento médico estandarizados, lo que precisa la modificación urgente de los instrumentos de evaluación de tal forma que prioricen los tests de razonamiento clínico situacional junto a la resolución de problemas complejos y las competencias comunicacionales por encima de la memorización o del reconocimiento de patrones que estas mismas herramientas ya reproducen con éxito. La evaluación de los retos éticos concluye que la integridad académica, el sesgo algorítmico, la protección de datos, el riesgo de despersonalización del cuidado son preocupaciones interrelacionadas que requieren marcos regulatorios institucionales explícitos, programas estructurados de formación docente y un seguimiento de los productos generados por estos sistemas bajo la supervisión humana.

Se concluye en que la región latinoamericana en general y el contexto ecuatoriano en particular tienen una brecha específica en la producción de evidencias propias, de lineamientos institucionales y de programas de formación del profesorado en esta materia. Debe por tanto resultar prioritario impulsar líneas de investigación y líneas de política pública educativa local en lugar de capas de marcos de referencia procedentes de otros contextos. Estos hallazgos responden al objetivo planteado y al mismo tiempo ofrecen un marco conceptual integrador para apoyar la toma de decisiones pertinentes en la formación del profesional de la salud conforme a la incorporación creciente de la inteligencia artificial generativa.

REFERENCIAS BIBLIOGRÁFICAS

1. Gordon M, Daniel M, Ajiboye A, Uraiby H, Xu NY, Bartlett R, et al. A scoping review of artificial intelligence in medical education: BEME Guide No. 84. Med Teach. 2024 Apr 2;46(4):446–70. doi:10.1080/0142159X.2024.2314198
2. Masters K. Ethical use of Artificial Intelligence in Health Professions Education: AMEE Guide No. 158. Med Teach. 2023 Jun 3;45(6):574–84. doi:10.1080/0142159X.2023.2186203
3. Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare. 2023 Mar 19;11(6):887. doi:10.3390/healthcare11060887
4. Preiksaitis C, Rose C. Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review. JMIR Med Educ. 2023 Oct 20;9:e48785. doi:10.2196/48785
5. Hale J, Alexander S, Wright ST, Gilliland K. Generative AI in Undergraduate Medical Education: A Rapid Review. J Med Educ Curric Dev. 2024 Jan 1;11. doi:10.1177/23821205241266697
6. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. PLOS Digital Health. 2023 Feb 9;2(2):e0000198. doi:10.1371/journal.pdig.0000198
7. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment. JMIR Med Educ. 2023 Feb 8;9:e45312. doi:10.2196/45312

8. Ali R, Tang OY, Connolly ID, Fridley JS, Shin JH, Zadnik Sullivan PL, et al. Performance of ChatGPT, GPT-4, and Google Bard on a Neurosurgery Oral Boards Preparation Question Bank. Neurosurgery. 2023 Nov;93(5):1090–8. doi:10.1227/neu.0000000000002551
9. Abdulai A, Hung L. Will ChatGPT undermine ethical values in nursing education, research, and practice? Nurs Inq. 2023 Jul 26;30(3). doi:10.1111/nin.12556
10. Choi EPH, Lee JJ, Ho MH, Kwok JYY, Lok KYW. Chatting or cheating? The impacts of ChatGPT and other artificial intelligence language models on nurse education. Nurse Educ Today. 2023 Jun;125:105796. doi:10.1016/j.nedt.2023.105796
11. Harrer S. Attention is not all you need: the complicated case of ethically using large language models in healthcare and medicine. EBioMedicine. 2023 Apr;90:104512. doi:10.1016/j.ebiom.2023.104512
12. Schuur F, Rezazade Mehrizi MH, Ranschaert E. Training opportunities of artificial intelligence (AI) in radiology: a systematic review. Eur Radiol. 2021 Aug 15;31(8):6021–9. doi:10.1007/s00330-020-07621-y
13. Ferreira AL, Lipoff JB. The complex ethics of applying ChatGPT and language model artificial intelligence in dermatology. J Am Acad Dermatol. 2023 Oct;89(4):e157–8. doi:10.1016/j.jaad.2023.05.054
14. Betti Galasso BJ. Inteligencia artificial generativa y formación docente: Competencias emergentes en América Latina – una revisión de alcance. Actualidades Investigativas en Educación. 2025 Dec 3;26(1):1–43. doi:10.15517/z4mcs052
15. Mortlock R, Lucas C. Generative artificial intelligence (Gen-AI) in pharmacy education: Utilization and implications for academic integrity: A scoping review. Exploratory Research in Clinical and Social Pharmacy. 2024 Sep;15:100481. doi:10.1016/j.rcsop.2024.100481
16. Snoswell CL, Falconer N, Snoswell AJ. Pharmacist vs machine: Pharmacy services in the age of large language models. Research in Social and Administrative Pharmacy. 2023 Jun;19(6):843–4. doi:10.1016/j.sapharm.2023.03.006
17. Khamis N, Ungaretti T, Tackett S, Chen B. From AI Literacy to Leadership: Milestones for Faculty Development in Health Professions Education. Med Sci Educ. 2025 Jun 4;35(6):2681–6. doi:10.1007/s40670-025-02438-0
18. Lopez M. From Novelty to Necessity: Artificial Intelligence in Medical Education. J Med Educ Curric Dev. 2025 Aug 14;12. doi:10.1177/23821205251369656
19. Currie GM. Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? Semin Nucl Med. 2023 Sep;53(5):719–30. doi:10.1053/j.semnuclmed.2023.04.008
20. Patino GA, Amiel JM, Brown M, Lypson ML, Chan TM. The Promise and Perils of Artificial Intelligence in Health Professions Education Practice and Scholarship. Academic Medicine. 2024 May 1;99(5):477–81. doi:10.1097/ACM.0000000000005636
21. Ramírez Domínguez CD, Alvarenga Somoza G, Olivares Guzmán NE, Cárcamo Trinidad MM, Salamanca Reyes AG. Avances en el uso de inteligencia artificial en la educación médica latinoamericana. Alerta, Revista científica del Instituto Nacional de Salud. 2025 Jan 22;8(1):88–95. doi:10.5377/alerta.v8i1.19194
22. Nouri Kandany V, Fabián Vásquez L, Valdez P, Jiménez Arias J, Valdez Disla A. Predictores Individuales y Profesionales del Uso de Inteligencia Artificial en Medicina: Un Análisis Multivariado en América Latina. Revista Española de Educación Médica. 2026 Jan 5;7(1). doi:10.6018/edumed.685291

23. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International. 2024 Mar 3;61(2):228–39. doi:10.1080/14703297.2023.2190148
24. Garabet R, Mackey BP, Cross J, Weingarten M. ChatGPT-4 Performance on USMLE Step 1 Style Questions and Its Implications for Medical Education: A Comparative Study Across Systems and Disciplines. Med Sci Educ. 2023 Dec 27;34(1):145–52. doi:10.1007/s40670-023-01956-z
25. Margolis A. Inteligencia artificial en la educación médica continua: ¿aliada inevitable o simplemente una herramienta más? Investigación en Educación Médica. 2025 Oct 1;14(56):146–53. doi:10.22201/fm.20075057e.2025.56.25735