

Intervenciones educativas basadas en inteligencia artificial en la educación médica y odontológica de pregrado: una revisión de alcance

AI-Based Educational Interventions in Undergraduate Medical and Dental Education: A Scoping Review

<https://doi.org/10.47606/ACVEN/MV0323>

Carmen Inocencia Quintana del Solar ^{1*}

<https://orcid.org/0000-0003-4676-5028>

cquintanad@unmsm.edu.pe

Sara Luz Quevedo Garcia ³

<https://orcid.org/0009-0000-3292-0029>

sara.quevedo@upsjb.edu.pe

Anace Méndez Torres ⁵

<https://orcid.org/0000-0003-1927-5816>

amendes@uroosevelt.edu.pe

Guadalupe Badillo Márquez ²

<https://orcid.org/0000-0002-0364-9329>

badillo.marquez.g@bine.mx

Germán Vicente Garay Flores ⁴

<https://orcid.org/0000-0002-7118-6477>

ggarayf101@ucvvirtual.edu.pe

José Eugenio Chafloque Capuñay ⁶

<https://orcid.org/0009-0002-4642-1540>

jchafloque@unprg.edu.pe

Recibido: 04/12/2025

Aceptado: 16/03/2026

RESUMEN

Introducción: La inteligencia artificial (IA) ha adquirido creciente relevancia en la educación médica y odontológica de pregrado, aunque la evidencia empírica sobre intervenciones educativas basadas en IA permanece dispersa y con limitada explicitación pedagógica. **Objetivo:** El objetivo de esta revisión de alcance fue mapear las intervenciones educativas basadas en IA implementadas en programas de pregrado en medicina y odontología, así como los resultados de aprendizaje, marcos pedagógicos y competencias de alfabetización en IA reportados. **Materiales y Métodos:** Se realizó una revisión de alcance siguiendo el marco metodológico de Arksey y O'Malley, los refinamientos de Levac et al. y la guía PRISMA-ScR. La búsqueda se efectuó en Scopus, PubMed/MEDLINE, ERIC y ProQuest Education Database entre 2021 y marzo de 2026. De 815 registros identificados, 31 estudios cumplieron los criterios de inclusión. Veintitrés correspondieron a educación médica y ocho a educación odontológica. Las intervenciones se agruparon en seis categorías: modelos generativos y chatbots, simulación clínica, retroalimentación automatizada, apoyo diagnóstico, entrenamiento procedimental e integración curricular. **Resultados:** Los resultados sugieren que la IA puede contribuir al aprendizaje cuando se integra con mediación docente, práctica repetida y retroalimentación oportuna. No obstante, la evidencia disponible es heterogénea y no demuestra una superioridad uniforme frente a los métodos tradicionales. **Conclusión:** Se concluye que la integración educativa de la IA requiere diseños pedagógicos explícitos, evaluación rigurosa y el desarrollo de competencias de alfabetización en IA en estudiantes y docentes.

Palabras clave: inteligencia artificial; revisión de alcance; educación médica; educación dental; alfabetización en IA.

1. Universidad Nacional Mayor de San Marcos
2. Benemérito Ins Nor del Edo Gral Juan Chrisóstomo Bonilla
3. Universidad Privada San Juan Bautista
4. Universidad Cesar Vallejo
5. Universidad Franklin Roosevelt
6. Centro de Medicina Biológica Divina Esperanza

* Autor de correspondencia: cquintanad@unmsm.edu.pe

ABSTRACT

Introductionn: Artificial intelligence (AI) has acquired increasing relevance in undergraduate medical and dental education, although empirical evidence on AI-based educational interventions remains scattered and with limited pedagogical explicitness. **Objective:** The objective of this scoping review was to map AI-based educational interventions implemented in undergraduate programs in medicine and dentistry, as well as the reported learning outcomes, pedagogical frameworks, and AI literacy competencies. **Materials and methods:** A scoping review was carried out following the methodological framework of Arksey and O'Malley, the refinements of Levac et al. and the PRISMA-ScR guideline. The search was conducted in Scopus, PubMed/MEDLINE, ERIC, and ProQuest Education Database between 2021 and March 2026. Of 815 records identified, 31 studies met the inclusion criteria. Twenty-three corresponded to medical education and eight to dental education. The interventions were grouped into six categories: generative models and chatbots, clinical simulation, automated feedback, diagnostic support, procedural training, and curriculum integration. **Results:** The results suggest that AI can contribute to learning when integrated with teacher mediation, repeated practice, and timely feedback. However, the available evidence is heterogeneous and does not demonstrate a uniform superiority over traditional methods. **Conclusions:** It is concluded that the educational integration of AI requires explicit pedagogical designs, rigorous evaluation and the development of AI literacy competencies in students and teachers.

Keywords: artificial intelligence; scope review; medical education; dental education; AI literacy.

INTRODUCCIÓN

La inteligencia artificial (IA) está presente y más visible en la educación en ciencias de la salud, especialmente en la formación médica. Sus usos ya no se limitan a tareas aisladas, sino que se han expandido hacia ámbitos como la enseñanza, la evaluación, la simulación, el razonamiento clínico y la interpretación de información compleja. En una revisión de alcance amplia sobre IA en educación médica, Gordon et al. documentaron un crecimiento sostenido del área y mostraron que las aplicaciones reportadas abarcan procesos de admisión, docencia, *assessment* y *clinical reasoning* [1]. Esto sugiere que el debate actual ya no se centra únicamente en si la IA “funciona”, sino en qué implica su incorporación desde una perspectiva pedagógica y curricular. En una línea complementaria, Shaw et al. señalaron que, aunque la IA ofrece posibilidades en múltiples dominios educativos, la evidencia empírica existente sigue siendo heterogénea y aún demanda mayor consolidación metodológica para respaldar decisiones formativas con mayor solidez [2].

En ese escenario, el avance reciente de la IA generativa (IA-Gen) ha acelerado la discusión y ha ensanchado las formas posibles de interacción educativa en el pregrado. Herramientas conversacionales basadas en *large language models* (LLM) vienen incorporándose gradualmente como apoyo al estudio, en la creación de materiales, en procesos de retroalimentación, así como en simulación y práctica clínica estructurada. No obstante, su expansión también ha reactivado inquietudes importantes: validez, supervisión, sesgos, integridad académica y uso responsable. Boscardin et al. sostienen que la aparición de IA-Gen obliga a fortalecer la alfabetización en IA tanto en docentes como en estudiantes, porque la integración pedagógica no debería depender solo de la disponibilidad tecnológica, sino de la capacidad crítica para reconocer alcances, límites e implicancias en la formación profesional [3].

La evidencia secundaria reciente también sugiere que los efectos educativos de la IA no se distribuyen de manera uniforme. Li et al., en un metaanálisis de ensayos controlados aleatorizados, hallaron que la enseñanza basada en IA-Gen no mostró diferencias significativas globales en la adquisición de conocimiento teórico frente a métodos tradicionales; sin embargo, sí se reportaron ventajas en habilidades prácticas y en satisfacción estudiantil en ciertos contextos [4]. Este resultado es relevante porque desplaza el debate desde una narrativa de reemplazo tecnológico hacia una lectura más situada: el valor educativo de la IA parece depender del *outcome* evaluado, del tiempo de exposición y, sobre todo, de cómo se integra la herramienta dentro del diseño pedagógico. En educación dental, la incorporación de IA también ha crecido de forma visible, aunque con una base empírica todavía menos consolidada que la observada en medicina. El-Hakim et al., en una revisión de alcance sobre IA en educación dental, identificaron aplicaciones en aprendizaje, evaluación y entrenamiento diagnóstico, pero también resaltaron limitaciones asociadas a desafíos metodológicos, vacíos de implementación y una necesidad persistente de mayor claridad pedagógica [5]. Estos hallazgos resultan especialmente pertinentes para el ámbito odontológico, porque muestran que la expansión tecnológica no siempre viene acompañada de una madurez curricular equivalente ni de una explicitación educativa suficiente, lo que dificulta traducir la innovación en orientaciones formativas estables y transferibles.

Este panorama cobra especial relevancia en el nivel de pregrado. En esta etapa, el estudiantado no solo adquiere conocimientos y habilidades técnicas básicas; también



construye formas de razonamiento clínico, criterios diagnósticos, hábitos de aprendizaje y marcos éticos que sostendrán su desempeño profesional. Por ello, evaluar la incorporación de IA en la formación médica y dental de pregrado exige ir más allá de su novedad o eficiencia computacional: importa su contribución real al aprendizaje, su coherencia con los fines educativos y su potencial para promover una alfabetización crítica orientada al uso profesional futuro de estas herramientas. En esa línea, Boscardin et al. subrayan que la alfabetización en IA es una necesidad emergente, precisamente porque usar una herramienta de manera funcional no equivale a comprender críticamente sus sesgos, su confiabilidad o los riesgos asociados a su aplicación [3].

Pese al crecimiento de publicaciones sobre IA en educación en ciencias de la salud, persisten vacíos que justifican una revisión de alcance enfocada específicamente en intervenciones empíricas dirigidas a estudiantes de pregrado en medicina y odontología. Una parte considerable de la literatura disponible mezcla niveles formativos, combina estudios empíricos con textos de opinión o enfatiza el potencial transformador de la IA sin precisar con suficiente detalle qué intervenciones han sido realmente implementadas, qué resultados de aprendizaje o desempeño se han evaluado y qué fundamentos pedagógicos sostienen dichas experiencias. Gordon et al. ya advertían la necesidad de desarrollar síntesis más delimitadas dentro del campo [1], y revisiones recientes en medicina y odontología coinciden en que la integración tecnológica avanza más rápido que la consolidación metodológica y curricular que debería acompañarla [5,2].

A pesar del avance de investigaciones sobre inteligencia artificial en educación en ciencias de la salud, mantiene una limitada sistematización de las intervenciones educativas basadas en IA dirigidas específicamente al nivel de pregrado. En muchos casos, la literatura combina diferentes niveles formativos o se centra en análisis conceptuales, lo que dificulta identificar con claridad qué estrategias pedagógicas han sido implementadas, qué resultados educativos han sido evaluados y qué marcos formativos sustentan dichas intervenciones.

En este contexto, la pregunta que orientó la presente revisión fue la siguiente: ¿qué intervenciones educativas basadas en IA se han implementado en educación médica y dental de pregrado, y qué resultados de aprendizaje, marcos pedagógicos o competencias relacionadas con IA se reportan en los estudios disponibles? A partir de esta pregunta, el objetivo de esta revisión de alcance fue mapear la evidencia empírica disponible sobre



intervenciones educativas basadas en IA en educación médica y dental de pregrado, con énfasis en los resultados de aprendizaje reportados y en los marcos pedagógicos o competencias de alfabetización en IA asociados.

MATERIAL Y METODOS

Se ejecutó una revisión de alcance (*scoping review*) con el fin de cartografiar la evidencia empírica disponible sobre intervenciones educativas basadas en inteligencia artificial en la educación médica y dental de pregrado. El foco se situó en los resultados de aprendizaje reportados, así como en los marcos pedagógicos y/o competencias de alfabetización en IA vinculados a dichas intervenciones. La revisión se sustentó en el marco metodológico propuesto por Arksey y O'Malley [6], incorporó las precisiones planteadas por Levac et al. [7]. El protocolo no se registró formalmente para esta revisión; sin embargo, el proceso metodológico se definió previamente por el equipo investigador y se condujo siguiendo las recomendaciones PRISMA-ScR [8], con el propósito de garantizar un proceso transparente tanto en la ejecución como en el reporte.

La revisión fue guiada por una pregunta estructurada según el marco PCC (Población, Concepto y Contexto): ¿qué intervenciones educativas basadas en IA se han implementado en educación médica y dental de pregrado, y qué resultados de aprendizaje, marcos pedagógicos o competencias relacionadas con IA se reportan en los estudios disponibles?

En este marco, la población estuvo conformada por alumnos de pregrado de medicina y/o odontología. El concepto se definió como intervenciones educativas donde la IA ocupaba un rol central en el diseño formativo, ya fuese mediante tutoría, simulación, retroalimentación, apoyo diagnóstico, entrenamiento procedimental, evaluación o como contenido incorporado al currículo. El contexto incluyó cursos, asignaturas, rotaciones, simulaciones y entornos preclínicos, clínicos, virtuales, híbridos o institucionales dentro de programas formales de formación médica y dental.

Se incluyeron estudios empíricos originales publicados en revistas científicas, en inglés, español o portugués, entre 2021 y la ejecución final de la búsqueda. Los estudios debían evaluar intervenciones educativas basadas en IA dirigidas a estudiantes de pregrado de medicina u odontología y reportar al menos un resultado asociado al aprendizaje,

desempeño, razonamiento clínico, precisión diagnóstica, adquisición de habilidades o competencias formativas.

Se excluyeron revisiones, editoriales, cartas, comentarios, protocolos, resúmenes de congresos sin texto completo, capítulos de libro, tesis y documentos teóricos sin evidencia empírica. Asimismo, se excluyeron estudios centrados únicamente en percepciones, aceptabilidad o satisfacción cuando no incluían medidas de aprendizaje o desempeño; investigaciones donde la IA no constituía una intervención educativa propiamente dicha; y estudios realizados en poblaciones distintas al pregrado cuando no era posible separar los datos correspondientes a estudiantes de pregrado.

La búsqueda bibliográfica se realizó en Scopus, PubMed/MEDLINE, ERIC y ProQuest Education Database. Para cada base se aplicó una estrategia principal y una estrategia suplementaria, adaptadas a su sintaxis particular. Posteriormente, ambas estrategias se integraron, unificaron y depuraron antes de iniciar el proceso de selección.

Como parte del control operativo, se ejecutó una búsqueda piloto el 1 de marzo de 2026, orientada a ajustar la estrategia y verificar la sensibilidad de las ecuaciones. La búsqueda final considerada para esta revisión se realizó el 3 de marzo de 2026.

Las estrategias combinaron términos vinculados a IA, educación médica y dental de pregrado, intervención educativa y resultados de aprendizaje. De manera complementaria, se incorporaron términos asociados a marcos pedagógicos y competencias de alfabetización en IA. No se incluyó literatura gris ni se realizó búsqueda manual secundaria como vía formal de incorporación de estudios.

Los registros recuperados se exportaron al gestor Mendeley para su organización y la eliminación de duplicados. Luego, los registros depurados fueron gestionado mediante hojas de cálculo en Microsoft en Excel para el cribado y para el registro sistemático de decisiones.

La selección se efectuó en dos etapas: (1) cribado por título y resumen y (2) evaluación de texto completo. En ambas fases, dos revisores evaluaron de forma independiente la elegibilidad conforme a criterios definidos previamente en el protocolo. Las discrepancias se resolvieron por consenso y, cuando persistieron, mediante la intervención de un tercer revisor.

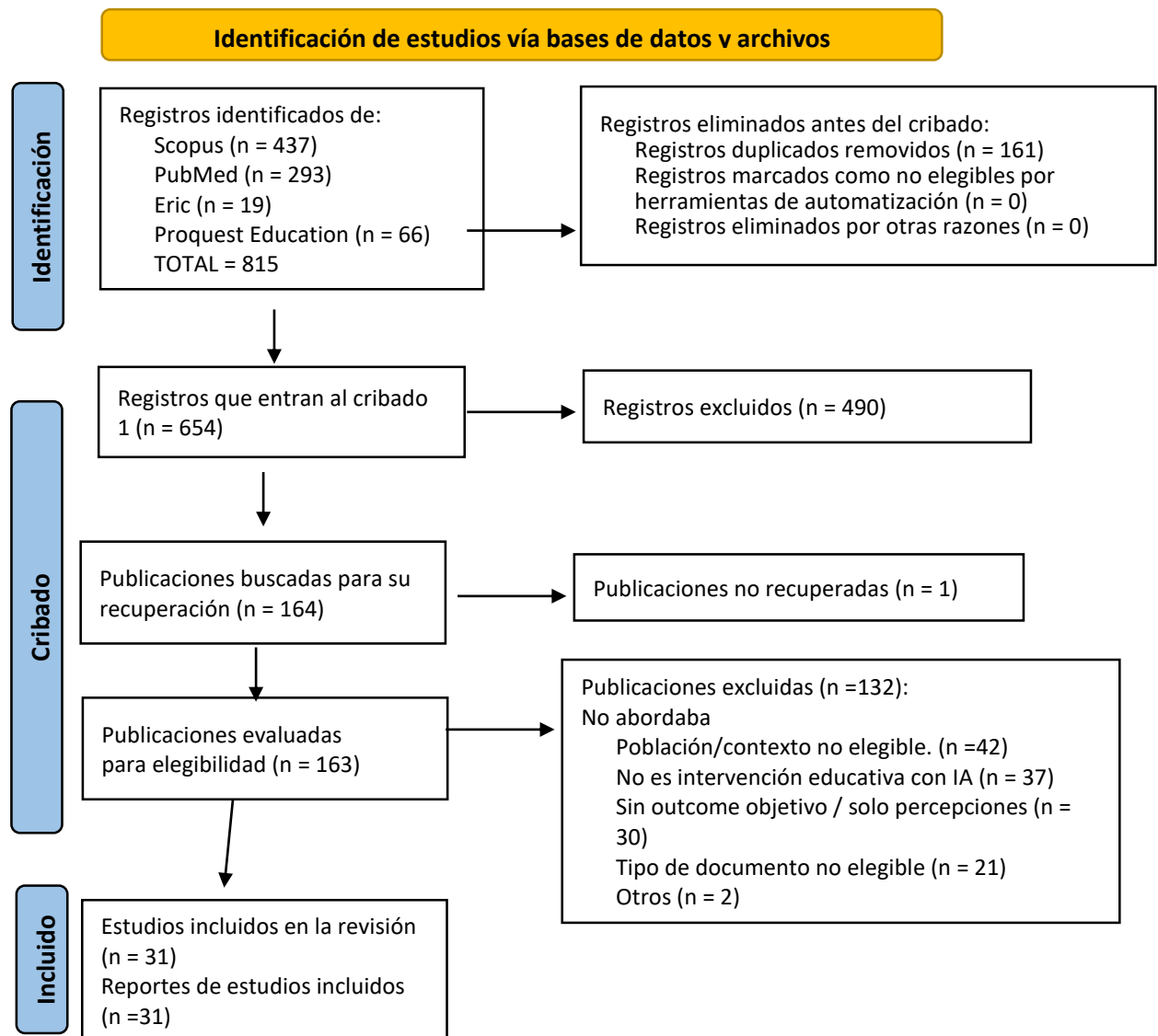
Se identificaron 815 registros en las cuatro bases de datos consultadas. Tras eliminar 161 duplicados, 654 registros avanzaron al cribado por título y resumen. Luego de esta etapa,

se buscaron 164 textos completos; uno no pudo recuperarse, por lo que 163 estudios fueron evaluados para elegibilidad. Finalmente, 31 estudios cumplieron los criterios de inclusión y fueron incorporados a la revisión.

Los motivos de exclusión en la fase de texto completo fueron registrados durante la etapa de elegibilidad y luego se agruparon para su reporte en el diagrama PRISMA en cinco categorías: (1) población o contexto no elegible; (2) intervención no centrada en IA educativa; (3) ausencia de outcomes objetivos de aprendizaje o desempeño; (4) tipo de documento no elegible; y (5) otros motivos metodológicos o de disponibilidad.

Figura 1.

Diagrama PRISMA del proceso metodológico de búsqueda, recuperación e inclusión de documentos para el estudio



La unidad de análisis correspondió a cada estudio incluido. La extracción incorporó variables bibliográficas, metodológicas y sustantivas, tales como: autor, año, país, disciplina, objetivo, diseño, muestra, nivel formativo, contexto educativo, tipo de intervención basada en IA, herramienta utilizada, comparador (cuando existió), duración, outcomes evaluados, instrumentos de medición y hallazgos principales.

De forma adicional, cuando la información estuvo disponible, se registraron el marco pedagógico reportado por los autores (cuando fue declarado explícitamente), las competencias de alfabetización en IA, las barreras y facilitadores de implementación, consideraciones éticas y limitaciones señaladas por los propios estudios.

Los hallazgos se integraron mediante un enfoque descriptivo y de mapeo, coherente con el propósito de una revisión de alcance. La evidencia se organizó en tablas de caracterización y en una síntesis narrativa orientada a describir los tipos de intervención educativa con IA, los contextos de implementación, los outcomes de aprendizaje reportados, las disciplinas involucradas y la presencia o ausencia de marcos pedagógicos y competencias de alfabetización en IA. No se realizó una evaluación formal de calidad metodológica ni metaanálisis, dado que el objetivo fue cartografiar el campo y no estimar efectos combinados.

RESULTADOS

Características generales de los estudios

Se incluyeron 31 estudios publicados entre 2022 y 2025 que evaluaron intervenciones educativas basadas en IA en pregrado de ciencias de la salud (ver Figura 1). La mayor concentración correspondió a 2024 y 2025, lo que sugiere un crecimiento reciente del campo más que una trayectoria consolidada. La descripción completa por estudio (país, diseño/muestra, contexto, intervención, comparador, outcomes y hallazgos) se presenta en la Tabla S1 del Anexo.

La mayoría de los estudios correspondió a educación médica ($n = 23$), mientras que ocho se desarrollaron en educación odontológica. En términos geográficos, predominaron estudios realizados en Europa, Asia y Norteamérica, sin identificarse estudios del corpus final procedentes de América Latina. En odontología, las intervenciones se concentraron principalmente en apoyo diagnóstico, interpretación de imágenes y entrenamiento preclínico. En medicina, en cambio, se observaron aplicaciones más diversas, incluyendo

simulación conversacional, entrenamiento procedimental, evaluación automatizada, ultrasonido y apoyo a actividades clínicas o académicas.

En cuanto al enfoque metodológico, se identificaron diseños experimentales y cuasiexperimentales, estudios de validación educativa y algunos enfoques mixtos. Una proporción importante utilizó comparadores: por ejemplo, enseñanza tradicional, condición sin IA u otra tecnología educativa, lo que permitió reportar outcomes objetivos asociados al aprendizaje o al desempeño. Esta heterogeneidad metodológica es relevante porque el corpus combina estudios orientados a medir efectividad, equivalencia, utilidad diagnóstica o factibilidad educativa, con diferentes niveles de control y de estandarización de los desenlaces.

Tipos de intervenciones educativas basadas en IA

A partir del mapeo del corpus, las intervenciones se agruparon en seis categorías: (1) modelos generativos y chatbots para aprendizaje teórico y apoyo docente; (2) simulación clínica y evaluativa, incluyendo pacientes simulados, OSCE y escenarios conversacionales; (3) retroalimentación personalizada o tutoría automatizada; (4) apoyo diagnóstico e interpretación de imágenes o mediciones; (5) entrenamiento procedimental y visualización anatómica; y (6) integración curricular y alfabetización en IA.

En términos funcionales, las categorías (1), (2), (3) y (5) se orientaron principalmente al andamiaje del aprendizaje mediante práctica guiada, feedback y entrenamiento repetible, mientras que la categoría (4) se enfocó en apoyo al juicio diagnóstico a través de reconocimiento de patrones, validación de hallazgos o estimaciones asistidas. La categoría (6) reunió las experiencias en las que la IA fue un objeto explícito de enseñanza o un componente formal del currículo. Esta organización muestra que la IA no operó como una innovación unitaria, sino como un conjunto de recursos con funciones pedagógicas distintas y con grados variables de integración curricula.

Tabla 1.

Mapa de tipos de intervención educativa basada en IA

Categoría de intervención	n (estudios)	Estudios (APA 7: autor-año)
Modelos generativos / chatbots (aprendizaje teórico / apoyo docente)	6	[9], [10], [11], [12], [13], [14]
Simulación (pacientes, OSCE, escenarios)	5	[15], [16], [17], [18], [19]
Retroalimentación / tutoría automatizada	3	[20], [21], [22]
Apoyo diagnóstico e interpretación (imágenes/mediciones)	9	[23], [24], [25], [26], [27], [28], [29], [30], [31]
Entrenamiento procedimental / visualización	5	[32], [33], [34], [35], [36]



Integración curricular / alfabetización en IA 3 [37], [38], [39]

Nota. El detalle por estudio (país, diseño/muestra, contexto, comparador, outcomes y hallazgos)

Resultados de aprendizaje y desempeño reportados

Los estudios reportaron *outcomes* predominantemente objetivos vinculados al aprendizaje o al desempeño académico y clínico; en varios casos se añadieron medidas secundarias como utilidad percibida, aceptabilidad o satisfacción. Los *outcomes* se agruparon en los siguientes dominios: conocimiento y rendimiento en pruebas; calidad de evaluación; desempeño en simulación y razonamiento clínico; precisión diagnóstica e interpretación; habilidades técnicas o procedimentales; y habilidades comunicacionales o de entrevista clínica.

De manera transversal, los resultados se reportaron con mayor consistencia cuando existió comparador y cuando los desenlaces fueron medidos con instrumentos o métricas estandarizadas, tales como rúbricas, puntajes de desempeño, precisión diagnóstica, concordancia o indicadores de ejecución (ver Tabla 2 y Tabla S1). En los estudios con medidas más estructuradas, la IA tendió a mostrar beneficios más nítidos en tareas delimitadas y observables, mientras que los resultados fueron más variables cuando las intervenciones dependieron de uso autónomo, objetivos amplios o marcos de implementación menos definidos.

Tabla 2.

Dominios de outcomes reportados

Dominio de outcome	n (estudios)*	Estudios (APA 7: autor-año)
Conocimiento y rendimiento en pruebas	11	[10], [15], [16], [11], [13], [21], [14], [17], [39], [18], [19]
Conocimiento y evaluación (calidad de ítems/psicometría)	3	[9], [11], [12]
Desempeño en simulación / razonamiento clínico	7	[15], [16], [25], [17], [35], [18], [19]
Precisión diagnóstica / interpretación	9	[23], [24], [25], [26], [27], [28], [29], [30], [31]
Habilidades técnicas / procedimentales	9	[24], [32], [33], [34], [20], [27], [28], [35], [36]
Comunicación / entrevista clínica	3	[22], [17], [18]
Otros outcomes	3	[37], [38], [31]

* **Nota.** Un mismo estudio puede contribuir a más de un dominio (la suma puede exceder 31).

Marcos pedagógicos y competencias de alfabetización en IA

En el corpus revisado, la explicitación de marcos pedagógicos fue heterogénea. Aunque la mayoría de las intervenciones definió objetivos educativos y evaluó outcomes, en numerosos estudios el sustento teórico-pedagógico no se reportó de forma explícita como



componente del diseño o de la evaluación (ver Tabla S1). Las excepciones más visibles correspondieron a experiencias de integración curricular o enseñanza híbrida, donde se identificaron referencias a modelos como *Community of Inquiry* Elhoshy et al., 2025, Kirkpatrick (Elhoshy et al., 2025), heutagogía (Digiacomio et al., 2025), *blended learning* (Höhne et al., 2025) o aprendizaje basado en escenarios (Zheng et al., 2025). Respecto a alfabetización en IA, se observaron dos patrones: primero, intervenciones donde la alfabetización en IA constituyó un objetivo formativo explícito, principalmente en experiencias de integración curricular; y segundo, intervenciones orientadas al uso funcional de herramientas basadas en IA para resolver tareas diagnósticas, comunicacionales o procedimentales, sin desarrollar necesariamente comprensión crítica sobre funcionamiento, limitaciones, sesgos o implicancias éticas. En conjunto, los hallazgos permiten distinguir entre propuestas para aprender con IA y propuestas para aprender sobre IA, lo que tiene consecuencias directas para el diseño curricular, la formación docente y las prioridades de investigación futura.

Barreras, facilitadores y consideraciones para la implementación

Aunque no todos los estudios reportaron de manera sistemática barreras y facilitadores, el corpus permitió identificar algunos patrones recurrentes. Entre los principales facilitadores destacaron la posibilidad de ofrecer retroalimentación inmediata y escalable, la repetición de la práctica en entornos estandarizados y la disponibilidad de métricas objetivas para seguimiento del desempeño. Estas ventajas fueron especialmente visibles en simulación clínica, entrenamiento procedimental, apoyo diagnóstico e intervenciones conversacionales con pacientes simulados.

En contraste, emergieron como barreras la necesidad de supervisión experta, la variabilidad en el rendimiento de algunas herramientas, la limitada explicitación pedagógica y la integración todavía fragmentaria de la IA dentro del currículo. De manera transversal, varios estudios sugirieron que la utilidad educativa de la IA depende menos de la novedad tecnológica que de la calidad del diseño instruccional, de la claridad de los criterios de desempeño y del grado de mediación docente con que la herramienta es implementada.

DISCUSIÓN

Los hallazgos de esta revisión de alcance muestran que la IA se está incorporando con rapidez en la educación médica y dental de pregrado, aunque la evidencia empírica disponible sigue siendo metodológicamente heterogénea y pedagógicamente desigual. En conjunto, los estudios sugieren que la IA puede aportar valor educativo cuando amplía oportunidades de práctica, ofrece retroalimentación oportuna y apoya tareas complejas, como simulación, razonamiento clínico o interpretación diagnóstica, especialmente en contextos donde la repetición, la estandarización y el *feedback* son relevantes para el aprendizaje [16,15,21,17]. Sin embargo, el beneficio reportado no parece depender solo de la sofisticación tecnológica, sino del alineamiento entre herramienta, objetivos de aprendizaje, mediación docente y contexto de uso [11,37,27].

Un patrón consistente es que la IA rinde mejor como recurso de apoyo y no como sustituto aislado del docente o del entorno formativo. Esto se aprecia con especial claridad cuando modelos generativos o *chatbots* se integran dentro de diseños híbridos, con clases, tutoría o actividades guiadas, en lugar de promoverse como uso autónomo sin estructura, donde los beneficios tienden a ser más variables [11,13,14]. De manera similar, las experiencias de integración curricular de contenidos sobre IA muestran resultados más sólidos cuando se implementan como propuestas organizadas, con secuencia didáctica y evaluación coherente con objetivos explícitos [37]. En otras palabras, la IA parece aportar más cuando fortalece un andamiaje pedagógico ya existente que cuando se adopta como innovación desanclada del currículo o como solución autosuficiente.

En intervenciones orientadas a retroalimentación personalizada, simulación clínica y entrenamiento procedimental, la inmediatez del *feedback* y la posibilidad de práctica repetida emergen como ventajas formativas frecuentes, asociadas a mejoras en desempeño o precisión [23,38,22,16,15]. En apoyo diagnóstico e interpretación de imágenes, con especial presencia en odontología y ecografía, varios estudios reportaron incrementos en exactitud o consistencia, lo que puede interpretarse como un andamiaje visual y cognitivo particularmente útil en tareas estructuradas, observables y susceptibles de retroalimentación inmediata [24,25,26,29,30,31]. Esta diferencia es importante porque sugiere que las herramientas de propósito específico tienden a mostrar beneficios más consistentes en *outcomes* delimitados, mientras que los modelos generativos de uso más

amplio parecen depender en mayor medida de mediación pedagógica y control contextual para producir ventajas claras.

El corpus también mostró divergencias que impiden sostener una narrativa uniforme de superioridad tecnológica. Algunos estudios informaron mejoras significativas frente a métodos tradicionales, mientras que otros reportaron resultados equivalentes, no inferiores o dependientes del *outcome* y del comparador, especialmente en entrenamientos técnicos o en escenarios comparados con realidad virtual y enseñanza *hands-on* [32,27,28]. Desde una perspectiva disciplinar, medicina presentó un repertorio más amplio de usos, incluyendo simulación, ultrasonido, entrenamiento técnico, evaluación y *chatbots* conversacionales, mientras que odontología concentró aplicaciones más consistentes en radiología, caries, cefalometría y apoyo diagnóstico [24,21,29,30]. Más que una diferencia de madurez entre disciplinas, este patrón probablemente refleja prioridades formativas distintas y oportunidades tecnológicas específicas según el tipo de tarea educativa predominante.

Un hallazgo crítico es la escasa explicitación de marcos pedagógicos en la mayoría de los estudios, lo que limita interpretabilidad, comparabilidad y replicabilidad. Esta omisión favorece lecturas donde el efecto educativo se atribuye casi exclusivamente a la herramienta, invisibilizando componentes centrales como secuenciación instruccional, criterios de desempeño, calidad del *feedback* y supervisión docente [37,11,19]. En la misma línea, los estudios que abordaron alfabetización en IA fueron todavía pocos: predominó aprender con IA por encima de aprender sobre IA, es decir, usar herramientas para tareas educativas sin desarrollar comprensión crítica sobre límites, sesgos, confiabilidad e implicancias profesionales [37,39]. Esta distinción es especialmente relevante porque formar usuarios eficaces no equivale a formar profesionales capaces de evaluar críticamente estas tecnologías en entornos clínicos y académicos.

En conjunto, esta revisión sugiere que la IA posee potencial educativo real en el pregrado médico y dental, pero su valor no es homogéneo ni automático. La pregunta clave deja de ser si la IA funciona y pasa a ser en qué condiciones, para qué propósitos y bajo qué marcos pedagógicos produce aprendizaje valioso, así como qué competencias críticas deben desarrollarse junto con su uso. Por ello, la incorporación curricular de la IA debería priorizar el alineamiento entre objetivos, evaluación y mediación docente, además de integrar consideraciones éticas y alfabetización en IA de manera explícita, para evitar

adopciones instrumentales que optimicen métricas inmediatas sin fortalecer competencias profesionales de largo plazo.

Asimismo, estos hallazgos tienen implicaciones relevantes para el diseño curricular en programas de ciencias de la salud, ya que sugieren la necesidad de integrar la inteligencia artificial no solo como herramienta tecnológica, sino como contenido formativo que fortalezca el razonamiento clínico, el pensamiento crítico y la toma de decisiones basadas en evidencia.

CONCLUSIÓN

Esta revisión de alcance mostró que la IA está adquiriendo una presencia creciente en la educación médica y dental de pregrado, especialmente en simulación clínica, apoyo diagnóstico e interpretación, retroalimentación automatizada, entrenamiento procedimental y uso de modelos generativos conversacionales. Los estudios incluidos sugieren que la IA puede contribuir al aprendizaje cuando se integra en experiencias formativas bien delimitadas, con objetivos explícitos, oportunidades de práctica repetida y mecanismos de retroalimentación oportunos, favoreciendo resultados en conocimiento, habilidades técnicas, razonamiento clínico y, en menor medida, competencias comunicacionales.

Sin embargo, la evidencia disponible no respalda una superioridad uniforme de la IA frente a los enfoques tradicionales. Los hallazgos variaron según el tipo de herramienta, el outcome evaluado, el diseño de implementación y el contexto educativo. En consecuencia, la principal conclusión de esta revisión no es que la IA reemplace o supere sistemáticamente las estrategias convencionales, sino que su valor educativo depende del grado de alineamiento entre tecnología, diseño pedagógico, evaluación y mediación docente.

Asimismo, la revisión identificó una brecha importante entre la velocidad de adopción tecnológica y la explicitación de fundamentos pedagógicos y competencias de alfabetización en IA. En la mayoría de los estudios, los marcos teóricos no se reportaron de manera clara, lo que limita la interpretabilidad, la comparación entre intervenciones y la replicabilidad de los hallazgos. De manera similar, fueron pocos los trabajos que abordaron de forma directa la alfabetización en IA, entendida como la comprensión crítica

de sus límites, sesgos, confiabilidad e implicancias éticas, pese a su relevancia para la formación profesional futura.

Concluimos que, la IA representa una oportunidad prometedora para enriquecer la educación médica y dental de pregrado, pero su incorporación no debería responder únicamente al impulso de la innovación tecnológica. Su integración curricular requiere decisiones deliberadas, evaluables y pedagógicamente justificadas, que articulen objetivos de aprendizaje, estrategias docentes, criterios de evaluación y formación crítica sobre el uso de estas herramientas. A partir de ello, futuras investigaciones deberían emplear diseños más robustos, reportar con mayor claridad el sustento pedagógico de las intervenciones, incorporar evaluación longitudinal de resultados y considerar de manera sistemática dimensiones éticas y de alfabetización en IA, con el fin de avanzar hacia una integración formativa, responsable y sostenible.

Esta revisión presenta limitaciones, siendo el principal que se restringió a estudios publicados en bases seleccionadas y en idiomas específicos, por lo que es posible que exista evidencia relevante no recuperada, incluyendo literatura gris o publicaciones fuera del alcance de las fuentes consultadas. Además, el corpus mostró heterogeneidad importante en diseños, contextos, herramientas y outcomes, lo que limita comparaciones directas entre intervenciones y no permite estimar efectos combinados. Asimismo, al tratarse de una revisión de alcance, no se realizó evaluación formal de calidad metodológica, por lo que los hallazgos deben interpretarse como un mapeo descriptivo del campo y no como evidencia definitiva de efectividad. Finalmente, varios estudios reportaron de manera limitada marcos pedagógicos, detalles de implementación y consideraciones éticas, lo que restringe la interpretabilidad y la posibilidad de replicación.

REFERENCIAS

1. Gordon M, Daniel M, Ajiboye A, Uraiby H, Xu NY, Bartlett R, Hanson J, Haas M, Spadafore M, Grafton-Clarke C, Gasiea RY, Michie C, Corral J, Kwan B, Dolmans D, Thammasitboon S. A scoping review of artificial intelligence in medical education: BEME Guide No. 84. *Med Teach*. 2024;46(4):446-470. doi:10.1080/0142159X.2024.2314198.
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*. 2025;35:1803-1816. doi:10.1007/s40670-025-02373-0.
3. Boscardin CK, Gin B, Black Golde P, Hauer KE. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. *Acad Med*. 2024;99(1):22-27. doi:10.1097/ACM.0000000000005439.

4. Li J, Yin K, Jiang X, Chen D. Effectiveness of generative artificial intelligence-based teaching versus traditional teaching methods in medical education: a meta-analysis of randomized controlled trials. *BMC Med Educ.* 2025;25(1):1175. doi:10.1186/s12909-025-07750-2.
5. El-Hakim M, Anthonappa R, Fawzy A. Artificial intelligence in dental education: a scoping review of applications, challenges, and gaps. *Dent J.* 2025;13(9):384. doi:10.3390/dj13090384.
6. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol.* 2005;8(1):19-32. doi:10.1080/1364557032000119616.
7. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* 2010;5:69. doi:10.1186/1748-5908-5-69.
8. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* 2018;169(7):467-473. doi:10.7326/M18-0850.
9. Ahmed, A., Kerr, E. & O'Malley, A. Quality assurance and validity of AI-generated single best answer questions. *BMC Med Educ.* 2025. doi:10.1186/s12909-025-06881-w.
10. Al Kahf S, Roux B, Clerc S, Bassehila M, Lecomte A, Moncomble E, Alababan E, de Montmolin N, Jablon E, François E, Friedlander G, Badoual C, Meyer G, Roche N, Martin C, Planquette B. Chatbot-based serious games: a useful tool for training medical students? A randomized controlled trial. *PLoS One.* 2023;18(1):e0278673. doi:10.1371/journal.pone.0278673.
11. Digiacoimo, A., Orsini, A., Cicchetti, R., Spadano, L., De Santis, S., Di Sessa, L., Vitale, M., Di Nicola, M., Tamborino, F., Basconi, M., De Archangelis, R., Salzano, G., Dello Stritto, G., Lannutti, P., Schips, L., & Marchioni, M. ChatGPT vs traditional pedagogy: a comparative study in urological learning. *World J Urol.* 2025;43:286. doi:10.1007/s00345-025-05654-w.
12. Fodor GH, Tolnai J, Rárosi F, Nagy A, Peták F. Artificial intelligence-based chatbots improve the efficiency of course orientation among medical students: a cross-sectional study. *BMC Med Educ.* 2025;25:1547. doi:10.1186/s12909-025-08146-y.
13. Gan W, Ouyang J, Li H, Xue Z, Zhang Y, Dong Q, Huang J, Zheng X, Zhang Y. Integrating ChatGPT in orthopedic education for medical undergraduates: randomized controlled trial. *JMIR Med Educ.* 2024;10:e57037. doi:10.2196/57037.
14. Huang, Y., Xu, B. B., Wang, X. Y., Luo, Y. C., Teng, M. M., & Weng, X. Implementation and evaluation of an optimized surgical clerkship teaching model utilizing ChatGPT. *BMC Med Educ.* 2024;24:1540. doi:10.1186/s12909-024-06575-9.
15. Bentegeac R, et al. ECOSBot: a multicenter validation pilot study of a generative AI tool for OSCE-based nephrology training. *Clin Kidney J.* 2025;18(10):sfaf308. doi:10.1093/ckj/sfaf308.
16. Brügge, E., Ricchizzi, S., Arenbeck, M., Keller, M. N., Schur, L., Stummer, W., Holling, M., Lu, M. H., & Darici, D. Large language models improve clinical decision making of medical students through patient simulation and structured feedback. *BMC Med Educ.* 2024;24:1391. doi:10.1186/s12909-024-06399-7.
17. Jiang Y, Fu X, Wang J, Liu Q, Wang X, Liu P, Fu R, Shi J, Wu Y. Enhancing medical education with chatbots: a randomized controlled trial on standardized patients for colorectal cancer. *BMC Med Educ.* 2024;24(1):1511. doi:10.1186/s12909-024-06530-8.

18. Yamamoto A, Koda M, Ogawa H, Miyoshi T, Maeda Y, Otsuka F, Ino H. Enhancing medical interview skills through AI-simulated patient interactions: nonrandomized controlled trial. *JMIR Med Educ.* 2024;10:e58753. doi:10.2196/58753.
19. Zheng K, Shen Z, Chen Z, Che C, Zhu H. Application of AI-empowered scenario-based simulation teaching mode in cardiovascular disease education. *BMC Med Educ.* 2024;24(1):1003. doi:10.1186/s12909-024-05977-z.
20. Fazlollahi AM, Bakhaidar M, Alsayegh A, Yilmaz R, Winkler-Schwartz A, Mirchi N, Langleben I, Ledwos N, Sabbagh AJ, Bajunaid K, Harley JM, Del Maestro RF. Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. *JAMA Netw Open.* 2022;5(2):e2149008. doi:10.1001/jamanetworkopen.2021.49008.
21. Gökkurt Yilmaz BN, Ozbey F, Yilmaz BE. Effect of artificial intelligence-assisted personalized feedback on radiographic diagnostic performance of dental students: a controlled study. *BMC Med Educ.* 2025;25(1):1403. doi:10.1186/s12909-025-07875-4.
22. Hershberger PJ, Pei Y, Bricker DA, Crawford TN, Shivakumar A, Castle A, Conway K, Medaramitta R, Rehtin M, Wilson JF. Motivational interviewing skills practice enhanced with artificial intelligence: ReadMI. *BMC Med Educ.* 2024;24:1-8. doi:10.1186/s12909-024-05217-4.
23. Aronovitz N, Hazan I, Jedwab R, Ben Shitrit I, Quinn A, Wacht O, Fuchs L. The effect of real-time EF automatic tool on cardiac ultrasound performance among medical students. *PLoS One.* 2024;19:e0299461. doi:10.1371/journal.pone.0299461.
24. Ayan E, Bayraktar Y, Çelik Ç. Dental student application of artificial intelligence technology in detecting proximal caries lesions. *J Dent Educ.* 2024;88:490-500. doi:10.1002/jdd.13437.
25. Dadon Z, Orlev A, Butnaru A, Rosenmann D, Glikson M, Gottlieb S, Alpert EA. Empowering medical students: harnessing artificial intelligence for precision point-of-care echocardiography assessment of left ventricular ejection fraction. *Int J Clin Pract.* 2023;2023:5225872. doi:10.1155/2023/5225872.
26. Fadillah R, Rikmasari R, Akbar S, Setiawan AS. Enhancing diagnostic precision: a calibration study integrating artificial intelligence for dental caries detection in dentistry training. *J Int Dent Med Res.* 2024;17(3):1100-1108.
27. Höhne E, Bauer E, Bauer C, Schäfer V, Gotta J, Reschke P, Vogl T, Yel I, Weimer J, Wittek A, Recker F. A comparative bicentric study on ultrasound education for students: app- and AI-supported learning versus traditional hands-on instruction (AI-Teach study). *Acad Radiol.* 2025;32(8):4930-4938. doi:10.1016/j.acra.2025.04.024.
28. Lau, Y. H., Acharyya, S., Wee, C. W. L., Xu, H., Saclolo, R. P., Cao, K., & Fong, W. K. Effectiveness of traditional, artificial intelligence-assisted, and virtual reality training modalities for focused cardiac ultrasound skill acquisition. *Ultrasound J.* 2025;17:61. doi:10.1186/s13089-025-00469-7.
29. Lin J, Liao Z, Dai J, Wang M, Yu R, Yang H, Liu C. Digital and artificial intelligence-assisted cephalometric training effectively enhanced students' landmarking accuracy in preclinical orthodontic education. *BMC Oral Health.* 2025;25(1):623. doi:10.1186/s12903-025-05978-4.
30. Ramezanzade S, Dascalu TL, Bakhshandeh A, Uribe SE, Ibragimov B, Bjørndal L. The impact of training dental students to use an artificial intelligence-based platform for pulp exposure prediction prior to deep caries excavation: a proof-of-concept randomised controlled trial. *Int Endod J.* 2025 Oct 10. doi:10.1111/iej.70046. Online ahead of print.

31. Schropp L, Sørensen APS, Devlin H, Matzen LH. Use of artificial intelligence software in dental education: a study on assisted proximal caries assessment in bitewing radiographs. *Eur J Dent Educ.* 2024;28:490-496. doi:10.1111/eje.12973.
32. Bogar, P. Z., Virag, M., Bene, M., Hardi, P., Matuz, A., Schlegl, A. T., Toth, L., Molnar, F., Nagy, B., Rendeki, S., Berner-Juhos, K., Ferencz, A., Fischer, K., & Maroti, P. Validation of a novel, low-fidelity virtual reality simulator and an artificial intelligence assessment approach for peg transfer laparoscopic training. *Sci Rep.* 2024;14:16702. doi:10.1038/s41598-024-67435-6.
33. Choi J, Lee Y, Kang GH, Jang YS, Kim W, Choi HY, Kim JG. Educational suitability of new channel-type video-laryngoscope with AI-based glottis guidance system for novices wearing personal protective equipment. *Medicine (Baltimore).* 2022;101(9):e28905. doi:10.1097/MD.00000000000028890.
34. Davidovic, V., Giglio, B., Albeloushi, A., Alhaj, A. K., Alhantoobi, M., Saeedi, R., Deraiche, S., Yilmaz, R., Tee, T., Fazlollahi, A. M., Ha, M., Uthamacumaran, A., Balasubramaniam, N., Correa, J. A., & Del Maestro, R. F. Effect of artificial intelligence-augmented human instruction on feedback frequency and surgical performance during simulation training. *J Surg Educ.* 2025;82:103743. doi:10.1016/j.jsurg.2025.103743.
35. McCarrick CA, McEntee PD, Boland PA, Donnelly S, O'Meara Y, Heneghan H, Cahill RA. A randomized controlled trial of a deep language learning model-based simulation tool for undergraduate medical students in surgery. *J Surg Educ.* 2025;82:103629. doi:10.1016/j.jsurg.2025.103629.
36. Nakao E, Igeta M, Kobayashi N, Kumazu Y, Otani Y, Murakami M, Kohno S, Hojo Y, Nakamura T, Kurahashi Y, Ishida Y, Shinohara H. Effectiveness of artificial intelligence-based visualization for surgical anatomy education: a cluster quasirandomized controlled trial. *Surgery.* 2025;188:109723. doi:10.1016/j.surg.2025.109723.
37. Elhoshy H, Rashed S, Fouad S, Abouzeid E, Talaat W. Evaluation of an artificial intelligence course integration into the undergraduate medical curriculum in Egypt: a mixed-methods study. *Educ Med J.* 2025;17(4):45-64. doi:10.21315/eimj2025.17.4.4.
38. Fazlollahi AM, Yilmaz R, Winkler-Schwartz A, Mirchi N, Ledwos N, Bakhaidar M, Alsayegh A, Del Maestro RF. AI in surgical curriculum design and unintended outcomes for technical competencies in simulation training. *JAMA Netw Open.* 2023;6(9):e2334658. doi:10.1001/jamanetworkopen.2023.34658.
39. Rath A. Empowering future dentists: a comprehensive mixed-methods exploration of artificial intelligence in personalizing year 3 clinical dental practice education. *J Dent Educ.* 2026;90:218-226. doi:10.1002/jdd.13939.