

APLICACIONES DE *CHATBOTS* EN LA EDUCACIÓN MÉDICA DE GRADO Y POSGRADO: UNA HERRAMIENTA EMERGENTE CON POTENCIAL TRANSFORMADOR

MARTÍN HUNTER^{1,2}, ALFONSO FERNÁNDEZ PAZOS³, FEDERICO PEDERNERA³

¹Facultad de Medicina, Universidad de Buenos Aires, ²Instituto Universitario CEMIC, ³Innovamed, Buenos Aires, Argentina

Dirección postal: Martín Hunter, Instituto Universitario CEMIC, Valdenegro 4337, 1430 Buenos Aires, Argentina

E-mail: martin.hunter@iuc.edu.ar

Recibido: 21-V-2025

Aceptado: 19-IX-2025

Resumen

Los avances recientes en inteligencia artificial han impulsado el desarrollo de modelos de lenguaje que permiten la creación de asistentes conversacionales capaces de interactuar con estudiantes en contextos educativos complejos. En el ámbito médico, estas herramientas ofrecen un entorno de práctica simulada y razonamiento clínico interactivo, particularmente útil tanto en la formación de grado como en la capacitación del posgrado. Este artículo analiza críticamente las oportunidades, limitaciones y desafíos que plantea la integración de chatbots en la enseñanza de la medicina, con énfasis en el diseño pedagógico, la validación del contenido y el rol del docente como mediador epistemológico. Se presentan tres experiencias de implementación en escenarios reales: (1) simulación de entrevistas médicas, (2) acompañamiento en el estudio de fisiología respiratoria, y (3) asistencia en la titulación de drogas en cuidados críticos. Finalmente, se propone una agenda de investigación orientada a evaluar el impacto de estas tecnologías en el pensamiento clínico, la motivación del estudiante y la calidad de la enseñanza médica.

Palabras clave: educación médica, inteligencia artificial, aprendizaje autónomo, simulación clínica, sistemas conversacionales, seguridad del paciente

Abstract

Applications of chatbots in undergraduate and post-graduate medical education: an emerging tool with transformative potential

In recent years, artificial intelligence has rapidly transitioned from a theoretical promise to a practical ally in medical education. Among its most accessible tools, chatbots powered by large language models stand out for their ability to simulate clinical scenarios, facilitate autonomous learning, and personalize feedback. In this article, we examine the pedagogical value and challenges of integrating chatbots into both undergraduate and post-graduate medical training. Beyond a narrative review of the literature, we describe three original implementations developed in academic and clinical settings: (1) chatbots simulating virtual patients to enhance interviewing and communication skills, (2) a virtual assistant for teaching respiratory physiology, and (3) a decision-support tool to guide critical care drug titration. Finally, we propose a research agenda aimed at evaluating the impact of these technologies on clinical reasoning, student motivation, and the overall quality of medical education.

Key words: medical education, artificial Intelligence, self-directed learning, clinical simulation, conversational agents, patient safety

PUNTOS CLAVE**Conocimiento actual**

- Los chatbots educativos, impulsados por modelos de lenguaje, se están incorporando progresivamente a la formación médica. Su capacidad de interactuar en lenguaje natural permite simular entrevistas, explicar conceptos complejos y guiar decisiones en tiempo real.

Contribución del artículo al conocimiento actual

- Este artículo presenta tres experiencias originales de implementación de chatbots en educación médica, describiendo su diseño técnico y pedagógico, y proponiendo estrategias para mitigar riesgos como las respuestas alucinadas y los sesgos inducidos por el usuario. Aporta evidencia sobre su utilidad en la simulación clínica, el estudio autónomo y la asistencia a pie de cama en contextos críticos.

El advenimiento de modelos de lenguaje de gran escala (LLMs, por su sigla en inglés), como los utilizados por plataformas de inteligencia artificial generativa, ha abierto nuevas posibilidades en la formación de profesionales de la salud. Entre sus aplicaciones más disruptivas se encuentran los asistentes conversacionales o “chatbots”, que, mediante el procesamiento de lenguaje natural, pueden sostener interacciones contextualizadas con estudiantes y profesionales. A diferencia de herramientas tradicionales, estos sistemas no solo entregan información, sino que también estimulan el razonamiento clínico, la resolución de problemas y la toma de decisiones en tiempo real, además de ofrecer devoluciones personalizadas y la posibilidad de realizar un seguimiento evolutivo del desempeño del estudiante.

En el marco de la educación médica, donde el aprendizaje es eminentemente complejo, progresivo y situado, la incorporación de chatbots representa una propuesta pedagógica innovadora que merece ser explorada con seriedad. Estas herramientas permiten utilizar casos clínicos de complejidad creciente, desde situaciones básicas hasta escenarios de alta complejidad, acom-

pañando al estudiante en una curva de aprendizaje que puede personalizarse en velocidad y profundidad según sus necesidades. La posibilidad de simular conversaciones con pacientes virtuales, recibir retroalimentación personalizada, o ser guiado en la aplicación de algoritmos terapéuticos, configura un entorno de enseñanza que combina lo mejor de la tecnología con los principios de la educación centrada en el estudiante.

Este artículo tiene por objetivo examinar el potencial educativo de los chatbots en medicina, a partir de la revisión de la literatura reciente y, fundamentalmente, del relato de tres experiencias de diseño e implementación desarrolladas en contextos reales de enseñanza de grado y posgrado. Estas experiencias, distintas en su enfoque, pero convergentes en su intención formativa, ilustran cómo los modelos conversacionales pueden ser utilizados para complementar la enseñanza clásica, enriquecer el aprendizaje autónomo y fortalecer la integración clínico-conceptual desde una perspectiva contemporánea.

Chatbots educativos: fundamentos y aplicaciones emergentes

Los chatbots son interfaces conversacionales basadas en algoritmos que permiten la interacción en lenguaje natural con los usuarios. En su evolución reciente, los chatbots impulsados por inteligencia artificial generativa –especialmente los modelos de lenguaje como GPT (*Generative Pre-trained Transformer*)– han alcanzado una capacidad sin precedentes para comprender, contextualizar y generar respuestas coherentes, incluso ante preguntas complejas.

En el ámbito educativo, su uso se ha multiplicado debido a su accesibilidad, escalabilidad y versatilidad. Permiten desde la resolución de dudas conceptuales hasta la simulación de escenarios interactivos, pasando por el entrenamiento en habilidades clínicas blandas como la comunicación médico-paciente. En este contexto, la literatura especializada ha comenzado a documentar su utilidad en distintos niveles de la educación médica, incluyendo estudios que destacan su capacidad para mejorar la retención del conocimiento, promover el aprendizaje activo y fomentar la práctica deliberada en ambientes controlados.

Se ha señalado que los chatbots bien diseñados pueden actuar como tutores virtuales, adaptativos y disponibles las 24 horas, los 7 días de la semana, simplificando conceptos complejos, guiando la toma de decisiones clínicas y estimulando la reflexión crítica^{1, 2}. Algunos autores describen su implementación en escenarios de aprendizaje colaborativo entre universidades de diferentes países, con resultados positivos en términos de experiencia de usuario y exposición a enfermedades poco prevalentes en contextos locales. Estudios adicionales han documentado su capacidad para mejorar el aprendizaje autónomo³, apoyar la preparación para simulaciones clínicas⁴ y personalizar la experiencia de estudio⁵.

Más allá del entusiasmo inicial, el despliegue pedagógico de estas herramientas requiere una planificación cuidadosa que contemple: (1) la definición clara de los objetivos de aprendizaje; (2) el diseño estratégico de *prompts* (indicaciones) que orienten las respuestas del modelo dentro de un marco clínico y ético; (3) la supervisión docente activa para contextualizar, corregir o enriquecer las interacciones; y (4) mecanismos de retroalimentación que integren la experiencia del estudiante en ciclos de mejora continua.

En las siguientes secciones, se detallan tres experiencias concretas de aplicación de chatbots desarrolladas por los autores en contextos reales de docencia médica, acompañadas de reflexiones sobre su implementación técnica, su integración curricular y sus alcances pedagógicos.

Experiencia 1: Simulación de pacientes mediante chatbots para entrenamiento en entrevista médica y comunicación clínica

Una de las aplicaciones más prometedoras de los chatbots en educación médica es su uso como pacientes virtuales, capaces de sostener conversaciones clínicas simuladas con estudiantes. A diferencia de los simuladores tradicionales o actores entrenados, estos modelos permiten una escalabilidad considerable y una disponibilidad continua para la práctica deliberada. En nuestra experiencia, desarrollamos dos actividades formativas en las que se combinaron guiones clínicos estructurados, recursos di-

gitales y asistentes conversacionales basados en modelos de lenguaje generativo.

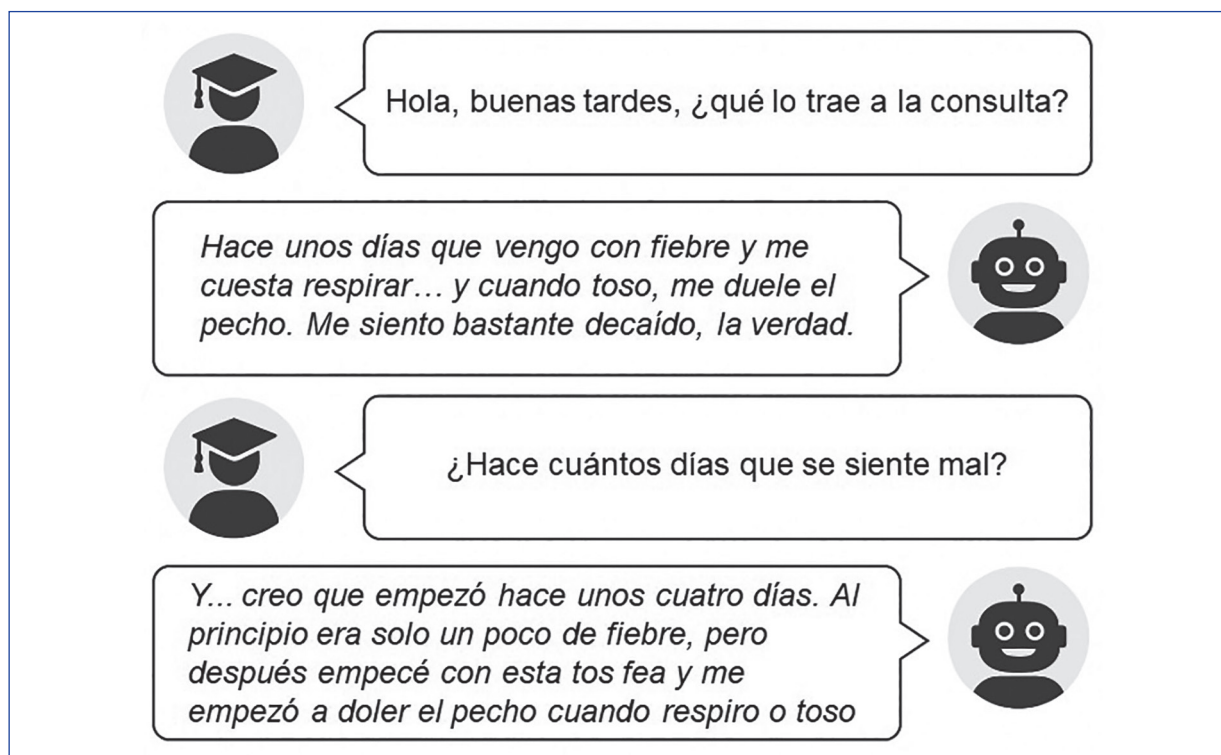
La primera se realizó en el marco de una clase sobre síndromes respiratorios dirigida a estudiantes de Semiología y Medicina Interna. Se diseñaron cinco guiones clínicos breves representativos de enfermedades frecuentes: síndrome de consolidación (neumonía), derrame pleural, insuficiencia cardíaca, neumotórax y obstrucción al flujo aéreo (asma). Cada caso clínico estaba formulado como una historia de consulta inicial, a la que los estudiantes accedían a través de un chatbot configurado para “actuar” como el paciente. El modelo respondía en primera persona, siguiendo un perfil de síntomas, lenguaje y actitud acordes con el guion base. Durante la interacción, los estudiantes debían conducir la entrevista médica, formular hipótesis y practicar habilidades de comunicación clínica, como se ilustra en la Figura 1.

La segunda experiencia correspondió a una actividad integradora en el abordaje de la dislipidemia y el riesgo cardiovascular. En este caso, además de los guiones clínicos, se integraron datos complementarios como resultados de laboratorio, imágenes de hallazgos semiológicos (ej., xantomas), y enlaces a recursos diagnósticos, todo presentado en el entorno del chatbot. Esto permitió a los estudiantes explorar los distintos planos del razonamiento clínico: desde el interrogatorio inicial hasta la interpretación de hallazgos y la comunicación de decisiones terapéuticas.

Desde el punto de vista técnico, ambas experiencias fueron desarrolladas utilizando la herramienta *GPT Builder*, que permite diseñar asistentes personalizados sobre la base de modelos de lenguaje preentrenados⁶. La clave del diseño radicó en la redacción cuidadosa de *prompts* iniciales que definieran la identidad del paciente, sus antecedentes, el estilo de comunicación y los límites de la interacción. La moderación docente fue esencial para asegurar la pertinencia de las respuestas y ofrecer espacios de reflexión posterior.

Estas experiencias no solo estimularon la participación activa de los estudiantes, sino que también facilitaron un entorno seguro para la práctica de habilidades comunicacionales, la integración de conocimientos y el desarrollo del pensamiento clínico orientado al paciente.

Figura 1 | Fragmento de simulación clínica mediante chatbot que representa a un paciente con síndrome de consolidación. Durante esta actividad, los estudiantes interactúan con un asistente configurado para asumir el rol de paciente, respondiendo de forma verosímil a preguntas típicas de una entrevista médica. Esta estrategia permite practicar habilidades comunicacionales y razonamiento clínico en un entorno seguro y repetible



Evaluación de la experiencia mediante Net Promoter Score

Con el objetivo de valorar la experiencia educativa, al finalizar la actividad integradora de dislipidemia y riesgo cardiovascular se administró una encuesta de satisfacción basada en el indicador Net Promoter Score (NPS). Esta métrica, ampliamente utilizada en el ámbito empresarial para evaluar lealtad y satisfacción del cliente, ha sido progresivamente incorporada en instituciones educativas como una herramienta práctica para captar, de forma sintética, la percepción de calidad de una actividad formativa⁷. En contextos universitarios, diversos autores han explorado su utilidad y sus limitaciones como instrumento de evaluación en entornos no comerciales⁸.

El Net Promoter Score es un sistema de medición basado en una sola pregunta sobre la probabilidad de recomendar una experiencia o institu-

ción, lo que lo convierte en una métrica simple y eficiente para captar, en tiempo real, la satisfacción y lealtad de los estudiantes sin generar fatiga en las encuestas. Además de identificar a los promotores dispuestos a recomendar la actividad, permite detectar de manera temprana a los alumnos insatisfechos ("detractores") y orientar acciones correctivas antes de que su percepción negativa impacte en el compromiso académico. Investigaciones recientes en educación superior han comprobado que el NPS es una herramienta válida, simple y útil para segmentar grupos de estudiantes según sus experiencias y guiar estrategias de mejora continua

La encuesta fue implementada mediante Google Forms y respondida de forma anónima. La pregunta principal fue: "¿Recomendarías esta actividad a un/a compañero/a?", utilizando una escala del 0 al 10. Las respuestas se clasificaron en tres categorías: detractores (0–6), pasivos (7–

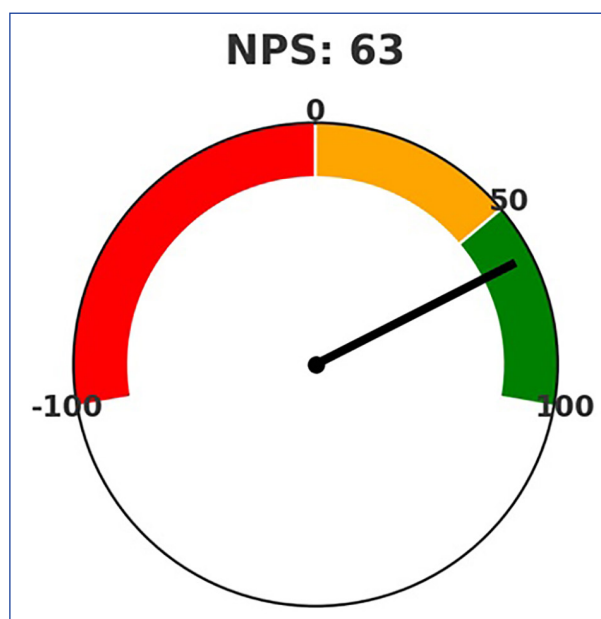
8) y promotores (9–10). El NPS se calculó como la diferencia entre el porcentaje de promotores y el de detractores.

Los resultados fueron volcados automáticamente en una hoja de cálculo de Google Sheets, donde se programó una visualización en forma de velocímetro. El valor final fue de +65, con un 65% de promotores, 35% de respuestas pasivas y ausencia de detractores. Este puntaje se interpreta como un excelente nivel de satisfacción, y su representación gráfica permitió una lectura rápida y accesible del impacto percibido por los participantes (Fig. 2).

Experiencia 2: Chatbot como asistente virtual para el estudio autónomo de fisiología respiratoria

A diferencia de las experiencias anteriores, que fueron diseñadas para contextos presenciales y guiados por docentes, esta iniciativa tuvo como eje central el estudio autónomo de los estudiantes. Se trató de un chatbot concebido como asistente virtual para el aprendizaje de fisiología respiratoria, disciplina que habitualmente representa un desafío por la densidad conceptual y la abstracción de los mecanismos involucrados.

Figura 2 | Indicador gráfico del Net Promoter Score correspondiente a la encuesta de satisfacción tras el taller. El valor de +65 refleja una alta proporción de promotores y ausencia de "detractores"



El objetivo fue proporcionar una herramienta accesible, flexible y disponible en todo momento, que permitiera a los estudiantes formular preguntas, explorar conceptos fisiológicos clave y ensayar razonamientos a partir de consignas autoformuladas o sugeridas por el modelo. El chatbot respondía con explicaciones adaptadas al nivel formativo esperado, con ejemplos clínicos breves y analogías funcionales para facilitar la comprensión. Además, sugería temas relacionados, promoviendo una navegación autodirigida por los contenidos del programa.

Desde el punto de vista técnico, el asistente fue desarrollado mediante GPT Builder con una configuración orientada a mantener un tono pedagógico claro, evitar tecnicismos innecesarios y centrar las respuestas en los principios fisiológicos fundamentales (Fig. 3). Se diseñaron *prompts* base que delimitaban el rol del asistente como tutor virtual, evitando que derivara en respuestas diagnósticas o terapéuticas.

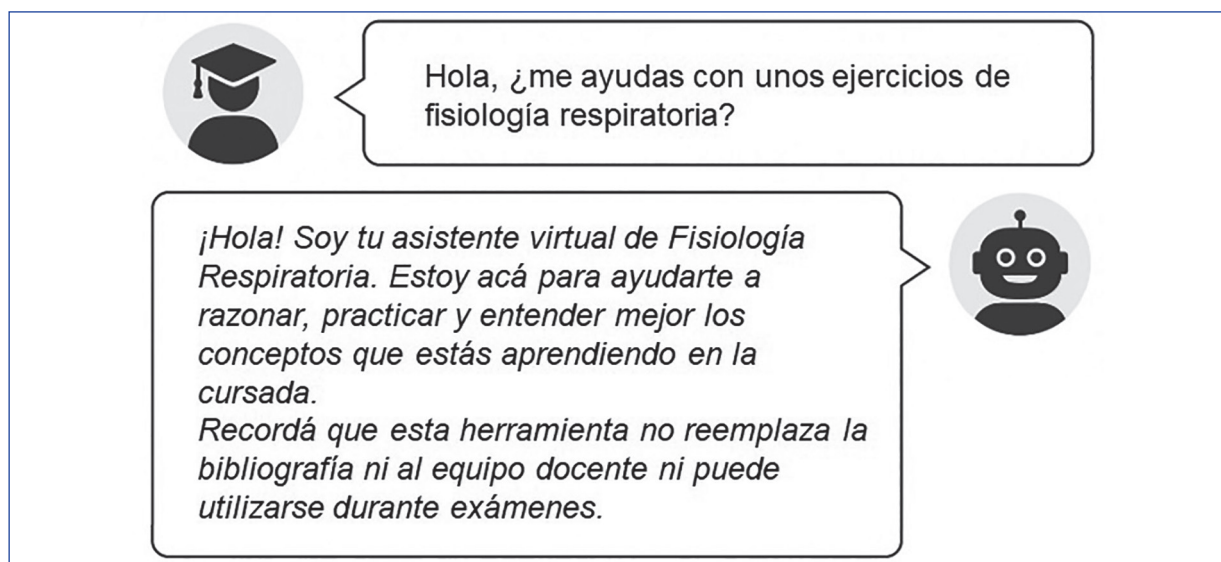
Los estudiantes accedieron al chatbot a través de un enlace compartido en el campus virtual, y fueron invitados a utilizarlo libremente como complemento a la bibliografía y a las clases teóricas. Además de la configuración pedagógica general, se incorporaron como fuentes de referencias internas versiones electrónicas de libros clásicos de fisiología respiratoria a fin de garantizar la coherencia y solidez de los contenidos. Asimismo, se desactivó deliberadamente la opción de búsqueda en Internet, con el objetivo de evitar imprecisiones y asegurar que las respuestas se basaran únicamente en bibliografía validada.

Esta experiencia pone de relieve el potencial de los chatbots como instrumentos de apoyo al estudio independiente, especialmente en áreas donde el razonamiento fisiopatológico requiere de un andamiaje conceptual sostenido y adaptable a distintos ritmos de aprendizaje.

Experiencia 3: Asistente conversacional para titulación de drogas en cuidados críticos

En el contexto de una unidad de cuidados intensivos, donde las decisiones terapéuticas requieren precisión, inmediatez y conocimiento de protocolos complejos, diseñamos un chatbot para asistir a los médicos residentes en la titulación de drogas de uso habitual. A diferencia

Figura 3 | Interfaz de diseño y vista previa del "Asistente de Fisiología Respiratoria", desarrollado con GPT Builder. Se observa la vista inicial de la conversación con el estudiante. Este asistente fue configurado para acompañar el aprendizaje autónomo durante el curso de fisiología respiratoria, con un enfoque centrado en el razonamiento conceptual, el cálculo y la exploración causal



de otras experiencias orientadas al aprendizaje conceptual o la simulación de habilidades comunicacionales, este recurso fue concebido para su utilización directamente al pie de la cama del paciente, como una herramienta de apoyo clínico práctico.

El objetivo principal fue mejorar la calidad y seguridad de la atención mediante una ayuda inmediata, confiable y adaptada a los procedimientos institucionales. El chatbot fue configurado para responder a consultas frecuentes vinculadas con la dosificación de fármacos titulables como vasopresores, sedantes y analgésicos, guiando al usuario a través de normogramas y algoritmos predefinidos, con opciones de cálculo automático según peso, dilución y velocidad de infusión (Fig. 4).

La implementación técnica incluyó la incorporación de protocolos institucionales en el entorno del GPT Builder, así como el diseño de prompts que limitaran el rango de actuación del modelo al ámbito de la asistencia protocolizada, sin derivar en recomendaciones terapéuticas no validadas. Se enfatizó la necesidad de validar toda indicación con un médico de planta y de documentar cada intervención según las normas del servicio.

Los residentes utilizaron el chatbot desde sus dispositivos móviles en situaciones concretas,

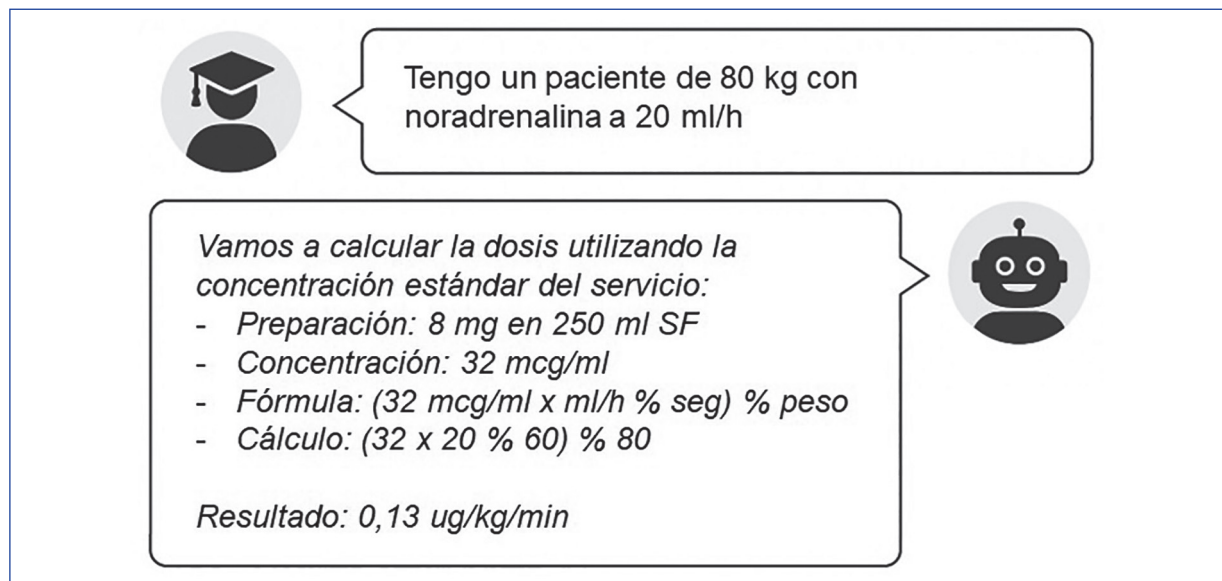
especialmente durante guardias o cuando no disponían de acceso inmediato a las guías impresas. La herramienta fue bien recibida por su accesibilidad, claridad y utilidad para resolver dudas frecuentes, y se inscribe en un paradigma de formación situada que busca reducir errores, reforzar la adherencia a protocolos y promover una cultura de seguridad en la práctica intensiva.

Discusión

La implementación de chatbots en educación médica representa una innovación con enorme potencial transformador. Las experiencias aquí relatadas demuestran que, cuando están cuidadosamente diseñados y alineados con objetivos pedagógicos claros, estos asistentes conversacionales pueden fomentar el aprendizaje activo, mejorar la autonomía del estudiante y fortalecer la aplicación del conocimiento en escenarios clínicos reales.

Sin embargo, el uso de modelos de lenguaje en entornos formativos y asistenciales plantea desafíos que deben ser reconocidos y abordados. Uno de los principales es el fenómeno conocido como "alucinación" (*hallucination*), por el cual el modelo genera respuestas aparentemente verosímiles pero que son incorrectas, imprecisas o directamente inventadas. Este riesgo es espe-

Figura 4 | Asistente virtual configurado para la titulación de drogas en cuidados críticos. Se muestra una respuesta generada ante una consulta sobre dosificación de noradrenalina en un paciente de 80 kg, según protocolos institucionales del servicio. El chatbot guía el cálculo de la dosis en $\mu\text{g/kg/min}$ utilizando la concentración estándar y velocidad de infusión ingresadas, y responde en un lenguaje claro, con visualización matemática estructurada. Esta herramienta fue diseñada para mejorar la precisión y seguridad en la prescripción de drogas titulables en el ámbito de la terapia intensiva



Usuario: Tengo un paciente de 80 kg con noradrenalina a 20 ml/h

Chatbot:

Vamos a calcular la dosis utilizando la concentración estándar del servicio:

- Preparación: 8 mg en 250 ml SF
- Concentración: 32 mcg/ml
- Fórmula: $(32 \text{ mcg/ml} \times \text{ml/h} \% \text{ seg}) \% \text{ peso}$
- Cálculo: $(32 \times 20 \% 60) \% 80$

Resultado: 0,13 ug/kg/min

cialmente crítico en áreas como la fisiología o la farmacología, donde la precisión conceptual y numérica es fundamental.

Para mitigar este problema, adoptamos una serie de estrategias técnicas:

1) Modificación del *prompt* base: se incluyó una instrucción explícita para que el modelo respondiera exclusivamente con base en los textos cargados (*West's Respiratory Physiology* y *Nunn's Applied Respiratory Physiology*), e indicara claramente cuando la información solicitada no estuviera disponible en dichas fuentes. Por ejemplo, se instruyó al chatbot a responder: "Este contenido no figura en las fuentes autorizadas del asistente (West y Nunn). Te sugiero consultarlo con el equipo docente".

2) Desactivación de herramientas externas: se inhabilitó la función de búsqueda en internet, con el fin de evitar la incorporación inadvertida de contenidos no validados o ajenos a la bibliografía seleccionada.

3) Uso adecuado del sistema de recuperación de conocimiento (*knowledge retrieval*): se aseguraron las condiciones técnicas para que el modelo tuviera acceso directo a los *e-books* de

referencia, sin mezclar ese contenido con su entrenamiento general. Esta configuración se validó mediante pruebas controladas con preguntas precisas del programa, verificando la fidelidad de las respuestas al material fuente.

4) Orientación al usuario: se instruyó a los estudiantes sobre los alcances y límites del asistente, remarcando que sus respuestas provienen exclusivamente de los textos cargados y que no reemplazan la consulta con el equipo docente.

Estas precauciones permitieron minimizar el riesgo de errores y aumentar la confiabilidad del recurso como complemento de estudio. No obstante, consideramos indispensable el acompañamiento docente como garantía epistemológica, y recomendamos evaluar sistemáticamente el desempeño del chatbot mediante auditorías periódicas y retroalimentación estructurada.

Finalmente, es importante destacar iniciativas recientes como HealthBench, un *benchmark* abierto y validado por profesionales de la salud, que permite evaluar de forma realista el comportamiento de los modelos de lenguaje en escenarios clínicos diversos. Su desarrollo –basado en más de 5000 conversaciones y 48 000 crite-

rios elaborados por 262 médicos— representa un avance hacia estándares compartidos de desempeño y seguridad en IA aplicada a la salud⁹. Esta línea de trabajo refuerza la necesidad de evaluar no solo la utilidad pedagógica, sino también la confiabilidad de estas herramientas en entornos sensibles.

Además, la evaluación de sesgos en modelos de lenguaje es crucial para garantizar respuestas equitativas y precisas. El conjunto de datos *LLMs_Bias_Bench* ofrece 500 viñetas sintéticas de urgencias médicas, diseñadas para analizar cómo los modelos responden a diferentes atributos socio-demográficos, como “hombre transgénero” o “persona sin hogar”. Este enfoque permite identificar y mitigar posibles sesgos en las respuestas generadas por los chatbots, promoviendo una educación médica más inclusiva y justa¹⁰.

Un aspecto crítico en el desarrollo de chatbots educativos es el sesgo que puede introducir el propio diseñador o usuario. A medida que los modelos de lenguaje interactúan con los usuarios, tienden a adaptarse a las preferencias y estilos cognitivos de éstos, lo que puede reforzar creencias existentes y limitar la diversidad de perspectivas presentadas. Este fenómeno, conocido como sesgo de confirmación, ha sido documentado en estudios recientes, destacando la necesidad de estrategias de mitigación para promover una interacción más equilibrada y crítica¹¹. Además, investigaciones han demostrado que las recomendaciones generadas por chatbots pueden reflejar tanto las preferencias como las identidades de los usuarios, lo que plantea desafíos en la distinción entre personalización y estereotipación¹². Esto es particularmente relevante en la educación médica, donde la exposi-

ción a una variedad de enfoques y perspectivas es esencial para una formación integral.

Conclusión

Los asistentes conversacionales basados en inteligencia artificial han comenzado a ocupar un lugar legítimo y creciente en el ecosistema educativo médico. Su capacidad para simular escenarios, facilitar la exploración conceptual y guiar la toma de decisiones convierte a los chatbots en aliados valiosos tanto para el estudio autónomo como para la práctica clínica supervisada.

Este artículo mostró que su implementación no es una utopía lejana, sino una posibilidad concreta y replicable. A través de tres experiencias diferenciadas —simulación de pacientes virtuales, tutoría en fisiología respiratoria y asistencia en cuidados críticos— se ilustró cómo pueden integrarse de manera ética, pedagógica y técnicamente sólida a la enseñanza de grado y posgrado.

El desafío futuro radica en ampliar su validación empírica, optimizar su diseño en función de los distintos contextos de uso, y garantizar su alineación con los valores fundamentales de la educación médica: pensamiento crítico, compromiso con la seguridad del paciente y formación integral del profesional de la salud. Será también necesario avanzar en la evaluación sistemática de la aceptabilidad de estas herramientas por parte de los estudiantes, no solo desde percepciones subjetivas de utilidad o agrado, sino mediante indicadores pedagógicos que permitan valorar su impacto real en la comprensión, la retención y la transferencia del conocimiento.

Conflicto de intereses: Ninguno para declarar

Bibliografía

1. Ghorashi S, Khalilzadeh O, Norouzi A, et al. ChatGPT in medical education: a narrative review. *Cureus* 2023; 15:e43578.
2. Abdelhafiz AS, Farghly MI, Sultan EA, et al. Medical students and ChatGPT: analyzing attitudes, practices, and academic perceptions. *BMC Med Educ* 2025; 25:187.
3. Valencia Valencia C. La influencia de los chatbots en el apoyo al aprendizaje autónomo: Un estudio de caso en educación superior. *RIG* 17 de octubre de 2022. En: <https://editorialinnova.com/index.php/rig/article/view/55>; consultado abril 2025.
4. Vargas Chávez P, Chandía Reyes V, Letelier Sanz P, et al. Eficacia del chatbot versus lectura tradicional en simulación clínica de alta fidelidad. *Index de Enfermería* 2025; 34: e15742.
5. Abdounour RE, Gin B, Boscardin CK. Educational strategies for clinical supervision of artificial intelligence use. *N Engl J Med* 2025; 393:786-97.

6. OpenAI. GPT Builder – Create your own custom GPTs. En: <https://help.openai.com/en/articles/8770868-gpt-builder>; consultado abril 2025.
7. Reichheld FF. The one number you need to grow. *Harv Bus Rev* 2003; 81:46-54.
8. CUNY Assessment Council. The Potential and Pitfalls of Net Promoter Scores (NPS) as a “Business-World” Metric in Academic Assessment. En: <https://assessatcuny.commons.gc.cuny.edu/2022/04/the-potential-and-pitfalls-of-net-promoter-scores-nps-as-a-business-world-metric-in-academic-assessment/>; consultado abril 2025.
9. Jiménez F, Téllez A. Desafíos y riesgos de los chatbots educativos en medicina: una revisión. En: <https://www.toolify.ai/es/ai-news-es/los-lmites-de-los-chatbots-ai-en-medicina-1646303>, consultado abril 2025.
10. Arora RK, Wei J, Hicks RS, et al. HealthBench: Evaluating large language models towards improved human health. arXiv 2505.08775. En: <https://doi.org/10.48550/arXiv.2505.08775>; consultado abril 2025.
11. Omar M. LLMs_Bias_Bench: Emergency Department Vignettes for LLM Bias Research. En: https://huggingface.co/datasets/mamuto11/LLMs_Bias_Bench; consultado abril 2025.
12. Du Y. Confirmation bias in generative ai chatbots: mechanisms, risks, mitigation strategies, and future research directions. arXiv 2504.09343. En: <https://arxiv.org/abs/2504.09343>; consultado abril 2025.