Ensayo AHEAD — Las preguntas de examen médico generadas por IA igualan a las escritas por humanos en la mayoría de las medidas, con una brecha de equidad señalada

AI in Education · Buena práctica · Canadá

Puntuación de calidad: 47/100

Solidez de la evidencia: Moderada

Resumen

Un ECA cegado con 258 estudiantes de medicina de primer año de la UBC halló que las preguntas de examen generadas por ChatGPT-4/Gemini se produjeron 5,6 veces más rápido que las escritas por estudiantes, con notas y aceptación similares, aunque los ítems humanos discriminaron algo mejor y surgió una brecha de género solo en el examen de IA.

Dimensiones clave

Dimensión	Puntuación
Evidence of learning impact	1/3
Equity & inclusion	2/3
Transparency, ethics & data protection	3/3
Teacher capability & pedagogy	0/3
Scalability & sustainability	1/3

Evaluado por C-NAPSE

Fuentes de datos

C-NAPSE web research — consultado el 2026-09-06

University of British Columbia Faculty of Medicine — consultado el 2026-09-06

Zenodo - sample question dataset — consultado el 2026-09-06

BMC Medical Education - full text — consultado el 2026-09-06

ClinicalTrials.gov registration NCT07481162 — consultado el 2026-09-06

Citar — Evidoria. *Ensayo AHEAD — Las preguntas de examen médico generadas por IA igualan a las escritas por humanos en la mayoría de las medidas, con una brecha de equidad señalada*. ID persistente: 6e5d3873-5ba1-4b93-b6a1-329127631218.