Essai AHEAD — Les questions d'examen médical générées par l'IA égalent celles rédigées par des humains sur la plupart des mesures, avec un écart d'équité signalé

AI in Education · Bonne pratique · Canada

Score de qualité : 47/100

Niveau de preuve : Modérées

Résumé

Un essai randomisé en aveugle mené auprès de 258 étudiants en médecine de première année à l'UBC a montré que les questions d'examen générées par ChatGPT-4/Gemini étaient produites 5,6 fois plus vite que celles rédigées par des étudiants, avec des notes et une acceptabilité similaires, bien que les items humains discriminent légèrement mieux et qu'un écart selon le genre soit apparu uniquement sur l'examen généré par l'IA.

Dimensions clés

Dimension	Score
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

Évalué par C-NAPSE

Sources de données

C-NAPSE web research — consulté le 2026-09-06

University of British Columbia Faculty of Medicine — consulté le 2026-09-06

Zenodo - sample question dataset — consulté le 2026-09-06

BMC Medical Education - full text — consulté le 2026-09-06

ClinicalTrials.gov registration NCT07481162 — consulté le 2026-09-06

Citer — Evidoria. *Essai AHEAD — Les questions d'examen médical générées par l'IA égalent celles rédigées par des humains sur la plupart des mesures, avec un écart d'équité signalé*. ID persistant : 6e5d3873-5ba1-4b93-b6a1-329127631218.