

Perceptions, acceptance and educational use of generative artificial intelligence among health professions students and educators: a rapid systematic review

Adesegun Nurudeen Osijirin

adesegunosijirin@fuahse.edu.ng

Federal University of Allied Health Sciences

Mohammed Sada Shamsudeen

Federal University of Allied Health Sciences

Felicia Nonye Egbeh

Federal University of Allied Health Sciences

Phina Chinelo Ezeagwu

Federal University of Allied Health Sciences

Deborah Ngozi Umah

Federal University of Allied Health Sciences

Charles Ifeanyi Anumaka

Federal University of Allied Health Sciences

Systematic Review

Keywords: generative artificial intelligence, ChatGPT, health professions education, medical education, nursing education, perception, acceptance, trust, rapid systematic review

Posted Date: August 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10534067/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Generative artificial intelligence (GenAI) is entering health professions education faster than curricula and governance can adapt. Educational adoption depends not only on technical capability but also on whether students and educators consider generated explanations, feedback and learning support useful, trustworthy and ethically acceptable.

Methods

A rapid systematic review was conducted and reported using PRISMA 2020 and rapid-review guidance. Structured searches of PubMed-indexed records, ERIC, Scopus-indexed publisher platforms and backward/forward citations covered 30 November 2022 to 29 July 2026. English-language, peer-reviewed empirical studies were eligible when they examined health professions students' or educators' perceptions, attitudes, trust, acceptance, concerns or educational use of GenAI. One reviewer screened, extracted and appraised all records, followed by a complete verification pass against source records. Methodological limitations were assessed with design-specific Mixed Methods Appraisal Tool criteria. Because measures and populations were heterogeneous, findings were synthesised narratively and thematically without meta-analysis.

Results

Sixty-five records were identified, 14 duplicates removed, 51 unique records screened and 30 full texts assessed. Twenty studies involving approximately 9,553 participants were included: 4 published in 2023, 7 in 2024, 8 in 2025 and 1 in 2026. Most were cross-sectional surveys and 16 focused predominantly on students. Acceptance was positive but conditional. Learners valued rapid explanations, summarisation, brainstorming, feedback, accessibility and productivity; educators recognised similar potential. Trust remained task-dependent and weakened when outputs affected clinical reasoning, grading or professional accountability. Recurrent concerns were hallucination and misinformation, plagiarism, over-reliance, erosion of critical thinking, privacy, bias and unclear institutional rules. Prior experience and AI literacy generally predicted more positive perceptions, whereas educators and clinicians tended to be more cautious than students. Common methodological limitations were convenience sampling, self-report measures, limited nonresponse analysis and rapidly ageing model-specific findings.

Conclusions

Health professions stakeholders do not simply accept or reject GenAI; they calibrate acceptance according to educational stakes, expected accuracy and human oversight. GenAI is best positioned as a supervised learning aid rather than an autonomous authority. Curricula should combine AI literacy, verification skills, explicit assessment rules and educator accountability. More longitudinal, experimental and low- and middle-income-country research is required.

Background

Generative artificial intelligence (GenAI) systems can produce explanations, summaries, questions, feedback, cases, code and multimodal learning resources in response to natural-language prompts. Their speed and apparent fluency make them attractive for personalised learning and educator productivity, but the same features can conceal fabricated references, factual errors, bias and context-insensitive advice [1, 2]. In health professions education, these limitations are consequential because educational information can shape clinical reasoning, professional identity and future patient care.

The emerging evidence base indicates that willingness to use GenAI is not equivalent to trust in its outputs. A student may rely on ChatGPT for brainstorming while rejecting its clinical recommendation; an educator may accept AI-assisted question drafting while refusing AI-generated grading. Perception therefore depends on the educational task, stakes, user experience, transparency, disciplinary context and degree of expert verification. General technology-acceptance measures alone cannot fully explain these task-specific judgements.

Recent evidence syntheses have mapped artificial intelligence and GenAI applications in health professions education, including curriculum integration and measured learning outcomes [3–6]. However, broad reviews combine performance studies, commentaries and tool evaluations, while profession-specific reviews do not fully compare learner and educator perceptions across medicine, nursing, dentistry,

pharmacy and allied health. The literature has also expanded rapidly during 2025–2026, making an updated stakeholder-centred synthesis necessary.

This review therefore addresses the following questions: (1) how are health professions students and educators using GenAI for education; (2) what benefits and learning value do they perceive; (3) what shapes trust, acceptance and intended use; (4) what risks and ethical concerns recur; and (5) what implications follow for curriculum, assessment and governance?

Methods

Review design and reporting

This rapid systematic review followed the PRISMA 2020 statement [7], PRISMA-S principles for reporting literature searches [8] and methodological guidance for rapid reviews. The review was converted from an earlier protocol draft after the search and synthesis had begun; consequently, it was not prospectively registered. This deviation is reported transparently and is considered in the limitations.

Table 1
Eligibility criteria.

Domain	Included	Excluded
Population	Students, trainees, faculty or educators in medicine, nursing, midwifery, dentistry, pharmacy, public health or allied health.	General university samples without separately extractable health-professions data; practising professionals with no education role.
Phenomenon	GenAI or large language models used or considered for teaching, learning, feedback, assessment preparation, research training or academic work.	Non-generative AI only; clinical decision support without an educational component.
Outcomes	Perceptions, attitudes, acceptance, trust, intended use, experience, concerns, integrity judgements or perceived educational value.	Model accuracy/performance only; no stakeholder-reported outcome.
Study design	Peer-reviewed quantitative, qualitative or mixed-methods empirical research.	Reviews, editorials, commentaries, letters, protocols, conference abstracts and preprints.
Dates/language	30 November 2022–29 July 2026; English full text.	Outside date range or unavailable English report.

Information sources and search strategy

On 29 July 2026, structured searches were conducted for PubMed-indexed records and ERIC records, supplemented by searches of Scopus-indexed journal publisher platforms and backward/forward citation searching of eligible studies and recent reviews. The core query combined GenAI terms (ChatGPT, generative artificial intelligence, large language model, LLM) with health-professions terms (medical, nursing, dental, pharmacy, allied health, physiotherapy, health professions) and stakeholder/outcome terms (student, educator, faculty, perception, attitude, acceptance, trust, experience, concern, use). Searches covered records published from the public release of ChatGPT on 30 November 2022 through 29 July 2026. Full source-specific strings and the search ledger are provided in Additional file 1. Journals were checked against Scopus coverage information available at the search date; only records from Scopus-indexed journals were retained.

Study selection and data extraction

Records were deduplicated by title, DOI and PubMed identifier. Titles and abstracts were screened against the eligibility criteria, followed by full-text assessment. One reviewer completed screening and extraction because this was a rapid review. To reduce error, the entire included set and all full-text exclusions were rechecked in a separate verification pass against publisher or PubMed records. Extracted variables were citation, country, profession, stakeholder group, sample size, design, GenAI focus, awareness/use indicators, perceived benefits, trust or acceptance findings, concerns and author recommendations. No automated screening or generative model was used to make eligibility decisions.

Methodological quality appraisal

Included studies were appraised using the design-specific criteria of the 2018 Mixed Methods Appraisal Tool (MMAT) [9]. Consistent with MMAT guidance, numerical total scores were not calculated. Appraisal considered appropriateness of sampling, representativeness, measurement quality, nonresponse risk, analysis, qualitative coherence and integration of mixed-methods components. Quality concerns informed interpretation but were not used to exclude studies.

Synthesis

Substantial heterogeneity in populations, instruments, response scales and outcomes precluded meta-analysis. Study characteristics were tabulated, quantitative findings were summarised by direction and magnitude where comparable, and qualitative/open-text findings were coded using thematic synthesis [30]. The synthesis followed SWiM principles [31] by stating the grouping logic, standardising the direction of findings and explaining why statistical pooling was inappropriate. Themes were developed iteratively around use, perceived value, calibrated trust, integrity, human oversight and implementation conditions.

Results

Study selection

The searches produced 65 records. Fourteen duplicates were removed, leaving 51 records for title and abstract screening. Twenty-one were excluded at this stage. Thirty full-text reports were assessed; ten were excluded because they evaluated model or learning performance without an eligible stakeholder-perception outcome ($n = 5$), involved ineligible professional or general populations ($n = 3$), or provided no extractable perception outcome ($n = 2$). Twenty studies were included (Fig. 1).

Characteristics of included studies

The 20 studies represented approximately 9,553 participants; the total is approximate because one report described its sample as “nearly 400”. Four studies were published in 2023, seven in 2024, eight in 2025 and one in 2026. Sixteen primarily studied students, two studied educators, and two included both learners and educators/faculty. Medicine dominated the evidence base, although nursing, dentistry, pharmacy, emergency medical services and allied health were represented. Eleven studies came from Asia or the Middle East, three from North America, two from Europe, one from Israel, one from Ukraine and two were multinational or global. Eighteen studies centred on ChatGPT; two addressed GenAI/LLMs more broadly. Sixteen used cross-sectional quantitative designs, three used mixed methods and one was qualitative.

Table 2
Characteristics and principal findings of included studies.

Study	Country/region	Participants	Design	GenAI focus	Principal findings	Key appraisal concern
Alkhaaldi et al. (2023)	United Arab Emirates	Recently graduated medical students; n = 265	Cross-sectional online survey	ChatGPT/AI experience and anticipated educational/clinical use	Overall optimism despite limited formal AI exposure; 53.2% expected improved learning. Prior experience predicted more positive views of AI-assisted care.	Voluntary single-centre sample; clear analysis.
Tangadulrat et al. (2023)	Thailand	Medical students (n = 124) and physicians (n = 115); total n = 239	Cross-sectional survey	Clinical practice and medical education	Medical students were more positive than physicians about educational use (83.1% vs 62.0%); both groups requested expert verification and deeper clinical context.	Group comparison appropriate; convenience recruitment.
Park (2023)	South Korea	Second-year pre-medical students; n = 99	Course-based cross-sectional study	ChatGPT as a feedback tool in leadership/communication learning	Students used ChatGPT iteratively before and after group work and valued rapid feedback, while noting the need to question and refine responses.	Authentic course context; one cohort and self-report.
Sallam et al. (2023)	Jordan	Health students; n = 458 (96.5% undergraduates)	Cross-sectional validation study	Acceptance measured with TAME-ChatGPT	Only 23.8% had previously heard of ChatGPT at the early study date. Perceived usefulness, ease of use, risks and behavioural intention formed distinguishable acceptance domains.	Validated instrument; non-probability sample.
Araujo & Cruz-Correia (2024)	Portugal	Medical informatics postgraduate students; n = 25	Mixed-methods survey	ChatGPT as a learning and academic-writing tool	Participants valued idea generation, text restructuring and support for non-native English writers, but stressed verification and guided integration into teaching.	Rich open responses; very small specialised sample.

Study	Country/region	Participants	Design	GenAI focus	Principal findings	Key appraisal concern
Ganjavi et al. (2024)	United States	Medical students from ≥ 28 schools; n = 415	Prospective cross-sectional survey	Awareness and use of ChatGPT/LLMs	Awareness was 96% and 52% had used an LLM for coursework. Students supported training and policy while expressing uncertainty about accuracy and appropriate boundaries.	Multisite reach; voluntary social-media/listserv recruitment.
Abu Hammour et al. (2024)	Jordan	Medical students; n = 244	Cross-sectional survey	Perceived benefits, practices and concerns	86.9% believed students could benefit; summarisation was strongly valued, while inaccurate information, dependence and inappropriate use were prominent concerns.	Multi-domain questionnaire; cross-sectional self-report.
Alharbi et al. (2024)	Saudi Arabia	Pharmacy, nursing and emergency medical services students; n = 354	Cross-sectional survey	Attitudes, perceptions and obstacles	91.2% were familiar with ChatGPT; 61.3% perceived increased productivity and 84.7% a positive educational impact. Accuracy and ethical use remained barriers.	Prevalidated questionnaire; single university.
Cherrez-Ojeda et al. (2024)	Americas (multinational)	Healthcare students; n = 2,661	Cross-sectional multicountry survey	Beliefs, ethics, trust and educational usefulness	42.99% were unaware of ChatGPT. Ethical confidence was associated with viewing it as trustworthy and useful; participants requested formal education and safeguards.	Large diverse sample; uneven country representation.
Pan et al. (2024)	China	Undergraduate medical students; n = 1,718	Cross-sectional multicentre survey	Use, trust and attitudes toward ChatGPT-like LLMs	34.5% had heard of and used LLMs. Although 77.2% reported relatively high trust, only 7.8% expressed strong trust; most attitudes were neutral or cautiously positive.	Large sample; regional rather than national coverage.

Study	Country/region	Participants	Design	GenAI focus	Principal findings	Key appraisal concern
Uribe et al. (2024)	Worldwide	Dental educators; n = 428	Cross-sectional international survey	AI chatbots/LLMs in dental education	31% already used AI and 64% saw educational potential. Respondents anticipated benefits for knowledge, research and decisions but worried about reduced human interaction and absent guidelines.	Global reach; 28% response rate and regional imbalance.
Kazley et al. (2025)	United States	Students across 12 health-professions programmes; nearly 400	Cross-sectional survey	Acceptable use and academic misconduct	Students distinguished assistive study use from graded production: many accepted learning support outside assessment but viewed AI use in tests and papers as a threat to integrity.	Validated adapted survey; exact response denominator not reported in abstract.
Moskovich & Rozani (2025)	Israel	Medicine, nursing and allied-health students; n = 217	Convergent mixed-methods survey	ChatGPT in healthcare and education	86.2% were familiar and mean perception was positive (4.04/5). Benefits centred on information retrieval and summarisation; verification, ethics and critical-thinking risks were recurrent themes.	Cross-profession comparison and integrated qualitative data; convenience sample.
Abouammoh et al. (2025)	Saudi Arabia	Medical students (n = 6) and faculty (n = 6)	Qualitative interview study	Earliest experiences, perceived benefits and limitations	Participants described accessibility and efficiency alongside concerns about reliability, assessment misuse, dependence and lack of institutional direction.	Appropriate qualitative approach; small, early-adopter sample.
Hong et al. (2025)	South Korea	Nurse educators; n = 120	Cross-sectional survey with open responses	Perceptions and experiences of GenAI	Educators were broadly positive but engagement was moderate. Accuracy, ethics and practical guidance	Relevant educator sample; descriptive design and self-selection.

Study	Country/region	Participants	Design	GenAI focus	Principal findings	Key appraisal concern
					separated interest from confident classroom adoption.	
Abou Hashish et al. (2025)	Saudi Arabia	Nursing students (n = 240) and educators (n = 40); interviews with 35	Convergent mixed methods	Knowledge, attitudes, concerns and use	Knowledge was moderate and attitudes positive. Educators reported greater concern about plagiarism, over-reliance and data accuracy; experience was the strongest predictor of perceptions.	Validated measures and integration of strands; single college.
Hu et al. (2025)	China	Medical students across Sichuan; n = 1,133	Cross-sectional survey	Use and prospective academic adoption	62.9% had used ChatGPT; information searching and assignments were common. Most were positive or neutral, but misinformation and plagiarism dominated concerns.	Large regional sample; voluntary survey and self-reported behaviour.
Fußhöller et al. (2025)	Germany	Medical students; n = 84	Cross-sectional mixed-question survey	Usage, educational impact and concerns	76.2% used ChatGPT for medical education, especially during examinations. Convenience and explanation were valued, but reliability and dependency limited uncritical acceptance.	Detailed use context; small sample.
Lymar et al. (2025)	Ukraine	Medical students, interns and PhD candidates; n = 244	Cross-sectional survey	Academic misconduct and AI use	84% used AI academically. Views of integrity were divided (36% misconduct, 26% acceptable, 38% undecided); efficiency benefits coexisted with error, critical-thinking and dependence concerns.	Timely ethics focus; locally developed instrument and single institution.

Study	Country/region	Participants	Design	GenAI focus	Principal findings	Key appraisal concern
Xu et al. (2026)	United States	Medical students; n = 157	National cross-sectional survey	Experiences and anticipated clinical/educational value	73.9% said use improved their view of AI's medical potential and 86.6% believed access could make them more effective physicians; respondents still requested formal training and governance.	National recruitment through student networks; modest response count.

Methodological quality

All studies had clear questions and data relevant to their stated aims. Nevertheless, methodological confidence was limited by the early and rapidly evolving nature of the field. Seventeen of 20 studies used voluntary convenience, snowball or institution-based recruitment; only five adequately discussed response patterns or representativeness. Nine reported a validated, adapted or pretested measurement instrument, while the remainder relied mainly on investigator-developed items. Most quantitative analyses were appropriate for descriptive objectives, but causal interpretations were not possible. The qualitative and mixed-methods studies added contextual depth, although small samples and limited integration in one study reduced transferability. Item-level appraisal is provided in Additional file 1.

Thematic synthesis

1. Rapid adoption accompanied by uneven literacy

Awareness changed markedly across the short publication period. In the earliest Jordanian survey, only 23.8% had heard of ChatGPT [13], whereas later US, Saudi and Israeli studies reported familiarity above 86% [15, 17, 22]. Actual educational use remained lower and more variable: 22% in one Saudi study [17], 52% among US medical students [15], 62.9% in Sichuan [26], 76.2% in Germany [27] and 84% among Ukrainian medical learners [28]. This gap between awareness and competent use was repeatedly linked to limited formal instruction. Familiarity or prior experience generally predicted more favourable perceptions [10, 15, 25].

2. Perceived learning value and productivity

Across professions, stakeholders valued immediate explanations, summarisation, brainstorming, language support, feedback, information retrieval and exam preparation. In the UAE, 53.2% expected ChatGPT to improve learning [10]; 83.1% of Thai medical students were positive about educational use [11]; 61.3% of Saudi healthcare students anticipated higher productivity and 84.7% perceived a positive educational impact [17]. Mixed-methods studies described GenAI as a convenient first-pass tutor and writing assistant, especially for generating alternative explanations and supporting users writing in a second language [14, 22, 25]. Educators similarly identified potential for knowledge acquisition, research and learning-resource development [20, 24].

3. Trust was calibrated to task and stakes

Positive attitudes did not amount to unconditional trust. Chinese medical students frequently reported “relatively high” trust, but only 7.8% reported strong trust [19]. Students and physicians in Thailand accepted low-stakes support more readily than clinical guidance, and physicians were more cautious than students [11]. In health-professions programmes, learners distinguished studying and brainstorming from submitting AI-generated tests or papers [21]. Trust was highest when GenAI served as an adjunct, sources could be checked and an educator retained responsibility; it weakened for clinical reasoning, grading and professionally accountable decisions.

4. Accuracy, hallucination and verification

The most consistent barrier was uncertainty about factual accuracy. Participants described plausible but incorrect answers, absent evidence, fabricated references and shallow or decontextualised responses [11, 16, 22, 26, 27]. Verification was therefore not a peripheral

recommendation but the main condition for educational acceptance. Stakeholders wanted source-linked outputs, discipline-specific tools, educator review and explicit teaching on cross-checking against authoritative resources. The pattern supports a calibrated-trust model: fluency increases usability but cannot justify epistemic authority.

5. Academic integrity, over-reliance and critical thinking

Integrity judgements varied by task. Learners commonly accepted grammar support, explanation and private study while treating undisclosed generation of assessed work as misconduct [21, 28]. In Ukraine, 36% classified academic AI use as misconduct, 26% considered it acceptable and 38% were undecided [28]. Students and educators worried that convenient answers could reduce independent reasoning, clinical judgement and skill development. Educators expressed stronger concern than students about plagiarism, dependency and inaccurate data [25]. These findings favour assessment-specific rules over blanket prohibition.

6. Human oversight, policy and AI literacy

Every professional group called for clearer institutional guidance. Desired curriculum content included prompt design, source verification, bias and privacy, citation and disclosure, acceptable assessment use, and recognition of situations requiring human expertise. Dental and nurse educators were positive but cautious, citing lack of training and policy as barriers to confident adoption [20, 24]. The evidence consistently positioned GenAI as an augmentation technology: useful for drafting and formative support, but requiring educator accountability for formal learning resources and high-stakes decisions.

Table 3
Integrated themes and implications for health professions education.

Theme	Evidence pattern	Risk if unmanaged	Recommended response
Uneven AI literacy	Awareness rose rapidly, but formal training lagged behind use.	Confident misuse and unequal benefit.	Embed staged AI literacy across curricula.
Perceived learning value	Speed, explanations, summaries, feedback and language support were repeatedly valued.	Substitution for active learning.	Use for scaffolding, retrieval practice and formative work.
Calibrated trust	Acceptance decreased as clinical or assessment stakes increased.	Automation bias or inappropriate reliance.	Match oversight intensity to task risk.
Accuracy and verification	Hallucination and weak sourcing were the dominant barriers.	Mislearning and downstream patient-safety risk.	Require source checking and educator validation.
Integrity and dependency	Stakeholders distinguished assistance from generation of assessed work.	Plagiarism, skill erosion and ambiguous misconduct.	Publish task-specific disclosure and assessment rules.
Governance and human role	Educators wanted policies, training and retained accountability.	Fragmented practice and hidden AI use.	Human-in-the-loop quality assurance and audit trails.

Discussion

Principal findings

This review synthesised 20 empirical studies across health professions and found a stable cross-study pattern: stakeholders were neither uncritically enthusiastic nor categorically resistant. They valued GenAI as a fast, accessible and adaptive learning aid, yet conditioned trust on verification, task stakes and human oversight. This “positive but conditional” stance appeared across medicine, nursing, dentistry, pharmacy and allied health, despite substantial variation in awareness and use.

The evidence also separates adoption from trust. Frequent users were usually more positive, but high use did not remove concerns about misinformation, integrity or dependency. This distinction matters for technology-acceptance research: intention to use may reflect convenience even when users doubt the truth or appropriateness of specific outputs. Future measures should therefore assess content credibility, reliance and verification behaviour separately from usefulness and ease of use.

Comparison with existing reviews

The findings extend broad reviews that have catalogued GenAI applications and educational outcomes [3–6]. Those syntheses show substantial growth in content generation, tutoring and assessment support; the present review clarifies the human conditions governing

adoption. It also complements evidence that GenAI can improve selected learning outcomes while demonstrating that educational benefit does not resolve questions of trust, authorship or professional accountability. The prominence of accuracy and integrity concerns across the 2025–2026 studies suggests that governance has become a central implementation issue rather than an ancillary ethical discussion.

Implications for curriculum, assessment and policy

Health professions programmes should replace ad hoc exposure with a longitudinal AI-literacy curriculum. Foundational teaching should cover what GenAI is and why fluent output may be wrong; applied teaching should require source verification, comparison with guidelines and reflection on uncertainty; advanced teaching should address bias, privacy, documentation, disclosure and patient-safety consequences. AI literacy should be assessed through demonstrated verification behaviour rather than self-reported confidence alone.

Assessment policies should specify permitted use by task. Brainstorming, language editing and formative explanation may be allowed with disclosure, whereas generation of assessed clinical reasoning or unverified patient-facing content requires restriction or close supervision. Authentic oral defence, process evidence, reflective commentary and version histories can reduce ambiguity while preserving legitimate assistive use. Faculty remain accountable for institutionally supplied AI-generated materials.

A practical calibrated-acceptance framework follows from the synthesis. Low-stakes, reversible tasks may use GenAI with learner verification; moderate-stakes materials require educator review and source documentation; high-stakes assessment, clinical reasoning and patient-facing outputs require expert validation and should not be delegated autonomously. Institutions should document the model/version, prompt, sources, reviewer and revision date for formally deployed AI-generated content.

Strengths and limitations

Strengths include a current search through 29 July 2026, cross-profession coverage, inclusion of student and educator perspectives, transparent screening counts, item-level methodological appraisal and synthesis of both quantitative and qualitative evidence. Restricting the corpus to peer-reviewed journals indexed in Scopus supported source consistency.

Several limitations qualify the conclusions. This was a rapid, single-reviewer process; although all decisions were verified in a second pass, independent duplicate screening and extraction were not available. Search access was limited to PubMed-indexed records, ERIC, publisher platforms and citation searching rather than direct export from every planned subscription database. The review was not prospectively registered. English-language restriction may have excluded local evidence. Most included studies used voluntary cross-sectional surveys, measured self-reported attitudes, and investigated rapidly changing versions of ChatGPT. Medicine and Asian/Middle Eastern settings were over-represented, while Africa, Latin America, midwifery, public health and many allied-health disciplines remained under-studied. Finally, heterogeneity prevented pooled estimates.

Research priorities

Future research should use longitudinal and experimental designs in authentic courses; compare learners with educators using common validated measures; record model version and actual prompts; distinguish perceived usefulness from calibrated reliance; observe verification behaviour; and examine effects on clinical reasoning, critical thinking and professional identity. Multisite studies in low- and middle-income countries are especially needed. Research should also evaluate multimodal GenAI, disability access, multilingual use, privacy and the effectiveness of specific institutional policies.

Conclusions

Health professions students and educators generally perceive GenAI as useful, but acceptance is contingent rather than absolute. Speed, accessibility, explanation and productivity drive use; accuracy, integrity, over-reliance, privacy and unclear accountability limit trust. The safest educational model is supervised augmentation: learners should verify, educators should validate, and institutions should define acceptable use according to task stakes. Responsible integration therefore requires more than access to a tool—it requires AI literacy, explicit governance and preservation of human professional judgement.

Abbreviations

AI
artificial intelligence
GenAI
generative artificial intelligence

LLM
large language model
MMAT
Mixed Methods Appraisal Tool
PRISMA
Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SWiM
Synthesis Without Meta-analysis.

Declarations

Ethics approval and consent to participate

Not applicable. The review synthesised publicly available research and involved no primary data collection.

Consent for publication

Not applicable.

Availability of data and materials

The complete search ledger, full-text exclusion reasons, extracted study-characteristics table and item-level MMAT appraisal are included in Additional file 1.

Competing interests

The author declares no competing interests.

Funding

The author received no financial or material support for this review.

Author's contributions

ANO conceived the review, developed and executed the search, screened and appraised the evidence, conducted the synthesis, drafted the manuscript, and approved the submitted version. MSS, FNE, PCE, DNU, and CIA revised the manuscript.

Acknowledgements

The author acknowledges no external funding or institutional assistance. ChatGPT (OpenAI) was used for language editing, document structuring and tabulation support. Eligibility judgements, evidence verification, interpretation and final responsibility remained with the author.

References

1. Kasneci E, Seßler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn Individ Differ*. 2023;103:102274. doi:10.1016/j.lindif.2023.102274.
2. Chan CKY, Hu W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *Int J Educ Technol High Educ*. 2023;20:43. doi:10.1186/s41239-023-00411-8.
3. Gordon M, Daniel M, Ajiboye A, et al. 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.
4. Basil M, Ahmed W, Hajeomar R, Strawbridge J, Lynch M, Mukhalalati B, et al. A scoping review of the use of generative artificial intelligence tools in health profession education. *BMC Med Educ*. 2026;26:291. doi:10.1186/s12909-025-08527-3.
5. Feigerlova E, Hani H, Hothersall-Davies E. A systematic review of the impact of artificial intelligence on educational outcomes in health professions education. *BMC Med Educ*. 2025;25:129. doi:10.1186/s12909-025-06719-5.
6. Pham T, et al. The impact of generative AI on health professional education: a systematic review in the context of student learning. *Med Educ*. 2025. doi:10.1111/medu.15746.

7. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71.
8. Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev*. 2021;10:39. doi:10.1186/s13643-020-01542-z.
9. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Educ Inf*. 2018;34(4):285–291. doi:10.3233/EFI-180221.
10. Alkhaaldi SMI, Kassab CH, Dimassi Z, Alsoud LO, Al Fahim M, Al Hageh C, et al. Medical student experiences and perceptions of ChatGPT and artificial intelligence: cross-sectional study. *JMIR Med Educ*. 2023;9:e51302. doi:10.2196/51302.
11. Tangadulrat P, Sono S, Tangtrakulwanich B. Using ChatGPT for clinical practice and medical education: cross-sectional survey of medical students' and physicians' perceptions. *JMIR Med Educ*. 2023;9:e50658. doi:10.2196/50658.
12. Park J. Medical students' patterns of using ChatGPT as a feedback tool and perceptions of ChatGPT in a Leadership and Communication course in Korea: a cross-sectional study. *J Educ Eval Health Prof*. 2023;20:29. doi:10.3352/jeehp.2023.20.29.
13. Sallam M, Salim NA, Barakat M, Al-Mahzoum K, Al-Tammemi AB, Malaeb D, et al. Assessing health students' attitudes and usage of ChatGPT in Jordan: validation study. *JMIR Med Educ*. 2023;9:e48254. doi:10.2196/48254.
14. Araujo SM, Cruz-Correia R. Incorporating ChatGPT in medical informatics education: mixed methods study on student perceptions and experiential integration proposals. *JMIR Med Educ*. 2024;10:e51151. doi:10.2196/51151.
15. Ganjavi C, Eppler M, O'Brien D, Storino Ramacciotti L, Ghauri MS, Anderson I, et al. ChatGPT and large language models awareness and use: a prospective cross-sectional survey of U.S. medical students. *PLOS Digit Health*. 2024;3(9):e0000596. doi:10.1371/journal.pdig.0000596.
16. Abu Hammour A, Abu Hammour K, Alhamad H, Nassar R, El-Dahiyat F, Sawaqed M, et al. Exploring Jordanian medical students' perceptions and concerns about ChatGPT in medical education: a cross-sectional study. *J Pharm Policy Pract*. 2024;17(1):2429000. doi:10.1080/20523211.2024.2429000.
17. Alharbi MK, Syed W, Innab A, Al-Rawi MBA, Alsadoun A, Bashatah A. Healthcare students attitudes opinions perceptions and perceived obstacles regarding ChatGPT in Saudi Arabia: a survey-based cross-sectional study. *Sci Rep*. 2024;14:22800. doi:10.1038/s41598-024-73359-y.
18. Cherrez-Ojeda I, Gallardo-Bastidas JC, Robles-Velasco K, et al. Understanding health care students' perceptions, beliefs, and attitudes toward AI-powered language models: cross-sectional study. *JMIR Med Educ*. 2024;10:e51757. doi:10.2196/51757.
19. Pan G, Ni J, et al. A cross sectional investigation of ChatGPT-like large language models application among medical students in China. *BMC Med Educ*. 2024;24:908. doi:10.1186/s12909-024-05871-8.
20. Uribe SE, Maldupa I, Kavadella A, El Tantawi M, Chaurasia A, Fontana M, et al. Artificial intelligence chatbots and large language models in dental education: worldwide survey of educators. *Eur J Dent Educ*. 2024;28(4):865–876. doi:10.1111/eje.13009.
21. Kazley AS, Andresen C, Mund A, Blankenship C, Segal R. Is use of ChatGPT cheating? Students of health professions perceptions. *Med Teach*. 2025;47(5):894–898. doi:10.1080/0142159X.2024.2385667.
22. Moskovich L, Rozani V. Health profession students' perceptions of ChatGPT in healthcare and education: insights from a mixed-methods study. *BMC Med Educ*. 2025;25:98. doi:10.1186/s12909-025-06702-0.
23. Abouammoh N, et al. Perceptions and earliest experiences of medical students and faculty with ChatGPT in medical education. *JMIR Med Educ*. 2025;11:e63400. doi:10.2196/63400.
24. Hong M, Shin H, Kim SS, De Gagne JC. Nurse educators' perceptions and experiences of generative artificial intelligence: a cross-sectional study analysis. *Comput Inform Nurs*. 2025;43(7):e01273. doi:10.1097/CIN.0000000000001273.
25. Abou Hashish EA, Alsayed SA, Abdel Razek NMF. Embracing AI in academia: a mixed methods study of nursing students' and educators' perspectives on using ChatGPT. *PLoS One*. 2025;20(7):e0327981. doi:10.1371/journal.pone.0327981.
26. Hu N, et al. Status and perceptions of ChatGPT utilization among medical students: a survey-based study. *BMC Med Educ*. 2025;25:831. doi:10.1186/s12909-025-07438-7.
27. Fußhöller M, et al. Perceptions, usage, and educational impact of ChatGPT among medical students in Germany: cross-sectional mixed methods survey. *JMIR Form Res*. 2025;9:e81484. doi:10.2196/81484.
28. Lyman L, Kuchyn I, Bielka K, et al. Academic misconduct and artificial intelligence use by medical students, interns and PhD students in Ukraine: a cross-sectional study. *BMC Med Educ*. 2025;25:1496. doi:10.1186/s12909-025-08100-y.
29. Xu Y, et al. Medical student experiences with ChatGPT: national cross-sectional study. *JMIR Form Res*. 2026;10:e76838. doi:10.2196/76838.

30. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol. 2008;8:45. doi:10.1186/1471-2288-8-45.
31. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368:l6890. doi:10.1136/bmj.l6890.
32. Bramer WM, Giustini D, Kramer BMR, Anderson PF. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Syst Rev. 2017;6:245. doi:10.1186/s13643-017-0644-y.

Figures

PRISMA 2020 flow diagram

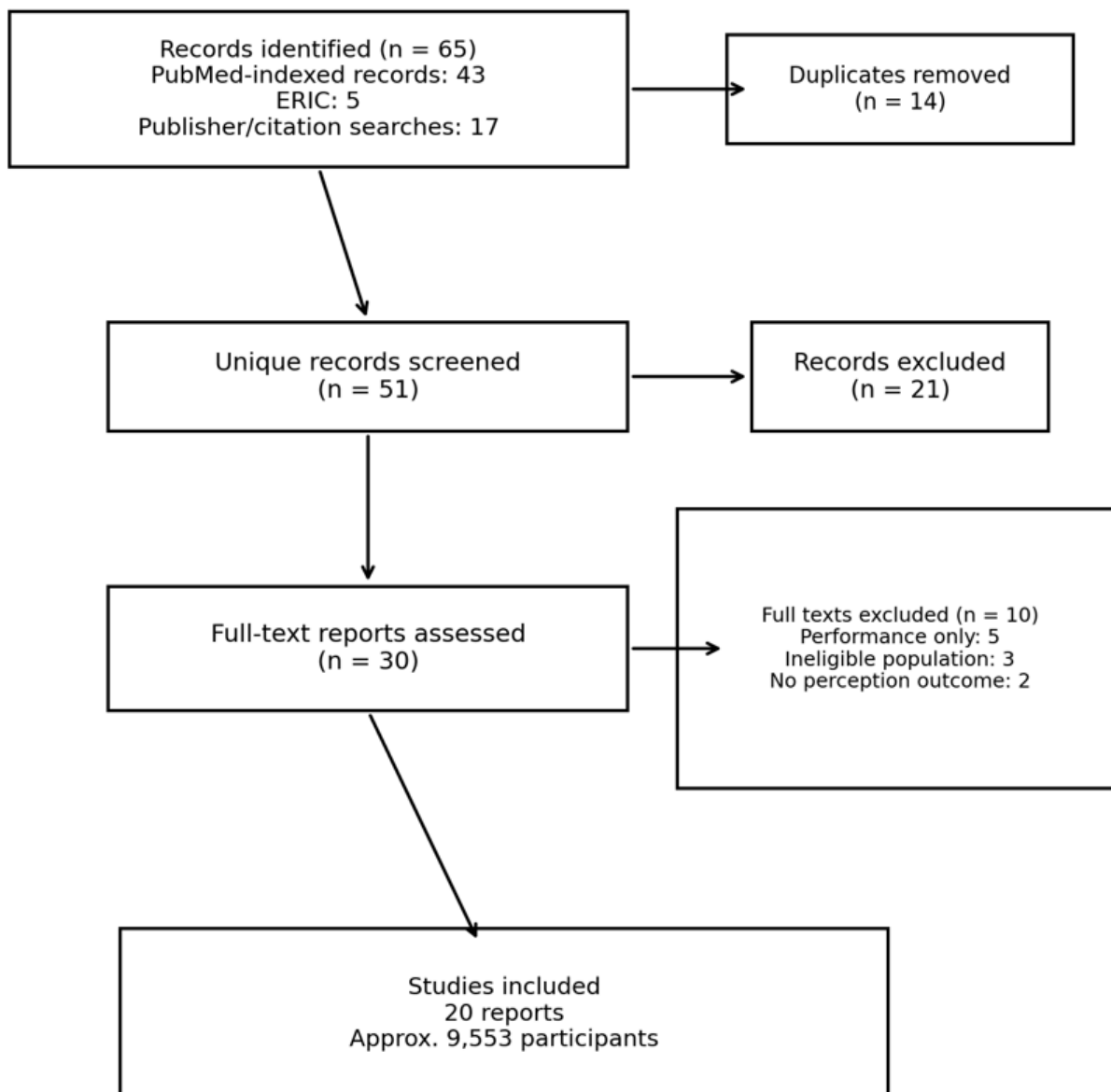


Figure 1

PRISMA 2020 flow diagram for study selection.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFile1SearchScreeningExtractionMMAT.docx](#)