ONLINE LEARNING
Volume 30 · Issue SI2 · 2026 · e5747
https://doi.org/10.24059/olj.v30iSI2.5747

The Integration of Generative Artificial Intelligence into Tertiary Health Education: A Scoping Review

Lok Tin Cheung
Haley Vu
Brooke Osborne
Nayana Parange
Sandhya Maranna
Adelaide University

Abstract

This scoping review aims to provide an overview of the integration of generative artificial intelligence (GenAI) in tertiary health education, including its practical applications in tertiary health education programs and how these applications are perceived by stakeholders. A literature search was conducted in five databases, including Ovid Medline, Ovid Emcare, Scopus, CINAHL, and ERIC. Articles integrating GenAI in tertiary health disciplines were included. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews checklist was followed in the reporting of results. Overall, 20 studies were included and three categories were identified: 1) application of GenAI in tertiary health education, 2) perceptions of GenAI use in health education, and 3) ethical and responsible use of GenAI in health education. Currently, GenAI is being integrated as a complementary tool to automate academic tasks and improve the critical thinking skills of tertiary health students through optimizing class designs. Although GenAI applications were well-received by students due to high accessibility, concerns over misinformation, over-reliance, and data confidentiality were consistently flagged. The review identified a lack of formal AI literacy training, highlighting the need for structured AI literacy training modules and clear ethical guidelines to support tertiary health students in using GenAI. GenAI applications are being rapidly integrated into tertiary health education, primarily for automated grading/feedback and improvement of learning outcomes. Future research is recommended to investigate the long-term impacts of GenAI and how to implement AI literacy training in current tertiary health education programs.

Keywords: Generative artificial intelligence, artificial intelligence, ChatGPT, tertiary health education, medical education, teaching and learning

Introduction

Generative artificial intelligence (GenAI) refers to a series of deep learning algorithms designed to generate new content from existing datasets through machine learning (Chan & Hu, 2023). At the core of GenAI are large language models, a type of artificial neural network trained to learn language structures and patterns from input datasets, and generate new, mostly text-based content based on prompts entered by users (Farazouli et al., 2023). The most prominent example of GenAI is the Chat Generative Pre-trained Transformer (ChatGPT), a GenAI tool that generates automatic responses to users' prompts in real time, based on patterns it learned from vast amount of training data (Gupta et al., 2023). Since its release in November 2022, GenAI technologies have undergone rapid developments, with new tools such as Stable Diffusion,

Corresponding author: Sandhya Maranna, sandy.maranna@adelaide.edu.au
This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).
© 2026 The Author(s)

Perplexity, and more advanced versions of ChatGPT developed and new features such as text-to-image conversion and advanced analysis (Shoja et al., 2023). With its growing capability to support interactive and iterative user engagement, GenAI has been increasingly applied across multiple sectors, including tertiary health education (Chan & Hu, 2023). One example of GenAI integration into tertiary health education includes the use of ChatGPT to generate multiple-choice questions for use in graduate medical examinations, effectively reducing the workload of university medical staff (Cheung et al., 2023). Another notable application involves employing ChatGPT to score the free-text and physical examination notes of medical students, displaying its potential as an efficient scoring and feedback tool in clinical education (Burke et al., 2024).

Tertiary Health Education and the Need for GenAI

Tertiary health education encompasses university-based education and training programs that prepare students for careers in the health professions, including medicine, nursing, midwifery, dentistry, pharmacy, allied health, and medical imaging (Allied health professions Australia, n.d; Gill & Galletly, 2025; Koontz & Podlog, 2025). Unlike many other higher education disciplines, health professions education integrates theoretical knowledge with clinical reasoning, psychomotor skill development, communication, professionalism, ethical decision-making, and supervised workplace learning. Educational approaches typically include simulation (Takhdat et al., 2026), Holderried problem-based learning, case-based discussions, laboratory training, and clinical placements, with graduates expected to demonstrate competence in providing safe, evidence-based, patient-centered care (Kotzki et al., 2026).

These distinctive characteristics make the integration of GenAI particularly important within health professions education. Healthcare systems are rapidly adopting AI-enabled technologies to support clinical decision-making, diagnostic interpretation, documentation, and patient management. Consequently, graduates require not only discipline-specific knowledge but also the ability to critically evaluate AI-generated information, understand its limitations, recognize potential biases, and use AI responsibly within professional and ethical boundaries (Daniel et al., 2025). Furthermore, GenAI offers opportunities to transform health education through personalised learning, adaptive tutoring, automated feedback, realistic clinical case generation, simulation, and support for evidence-based learning. At the same time, concerns regarding academic integrity, learner over-reliance, data privacy, equity (Atif et al., 2026), and clinical accountability necessitate careful educational integration supported by appropriate governance and curriculum design (Crompton et al., 2026).

Literature Review

Emerging research has begun to explore the role of GenAI in tertiary health education, with existing studies primarily looking at aspects such as technical performance, the perceived benefits, and potential drawbacks on student learning as seen in recent reviews by Preiksaitis & Rose, (2023) and Shoja et al., (2023). Furthermore, although current studies provide valuable insights on GenAI in tertiary health education, the evidence remains dispersed and often discipline-specific such as dentistry (Uribe et al., 2025), nursing (Hardie et al, 2026; Hinsche et al., 2026), and pharmacy (Mortlock & Lucas, 2024). This fragmentation makes it difficult for educators, curriculum designers, and policymakers to grasp the broader landscape of GenAI use, evaluate its pedagogical value, or anticipate risks emerging from its adoption. Given the rapid

expansion and diversification of GenAI technologies, a consolidated understanding of how these tools are being integrated into broad health education is needed.

Toohey et al. (2025) proposed a framework through their scoping review for a structured approach for integrating GenAI into health education, emphasizing responsible adoption across curriculum, assessment, governance, and workforce preparation. It highlights the need for AI literacy, educator capability, assessment design and evidence-informed implementation to support safe, effective, and sustainable GenAI-enhanced learning. Lin et al. (2025) in their review identified key challenges in integrating GenAI into medical education, including inaccurate or unreliable AI-generated content, academic integrity concerns, privacy and ethical risks, limited AI literacy among educators and students, over-reliance on AI, digital inequities, and insufficient evidence regarding educational effectiveness. The review highlights the need for governance, training, and responsible implementation strategies (Lin et al., 2025). A systematic review by Pham et al. (2025) investigated how students in health profession utilize GenAI using the Laurillard's six learning types and compared it to conventional and digital technologies. After evaluating 33 studies, they found that there was a shift to autonomous learning with GenAI and a reduction in collaborative learning (Pham et al., 2025).

Several scoping and systematic reviews have examined the role of GenAI in higher education; however, important gaps remain regarding its application within health professions education. Existing reviews of GenAI in higher education have primarily focused on broad educational contexts, including student engagement, academic integrity, assessment practices, and teaching efficiency, and have generated general recommendations for educators and institutions (Belkina et al., 2026; Daniel et al., 2025). While valuable, these reviews do not adequately address the distinctive characteristics of health professions education, which requires integration of theoretical knowledge with clinical reasoning, practical skill development, professional judgement, ethical decision-making, and patient-centered care.

Within health education, previous reviews have predominantly examined artificial intelligence more broadly rather than specifically focusing on GenAI and large language models (Kovalainen et al., 2025). These reviews provide important insights into the broader role of AI in healthcare education but do not capture the unique opportunities and challenges introduced by GenAI, including its capacity to generate clinical cases, provide conversational feedback, support personalised learning, assist with simulation-based education, and influence assessment practices. Furthermore, reviews focusing on specific health disciplines, such as medicine (Lin et al., 2025), provide valuable discipline-specific insights but may not represent the diversity of health professions education, including nursing, allied health, pharmacy, dentistry, and medical imaging.

Therefore, a gap exists for a comprehensive synthesis that specifically examines the integration of GenAI across tertiary health professions education. The current scoping review addresses this gap by mapping the rapidly emerging evidence base following the widespread adoption of ChatGPT and other large language models, identifying common themes across health disciplines, and highlighting discipline-specific opportunities and challenges.

The primary objective of this scoping review is to systematically explore the current landscape of how GenAI is currently being integrated into tertiary health education (i.e.,

undergraduate or postgraduate programs at university level) for health professional students (e.g., allied health, medical and dentistry cohorts). Specifically, the review aims to categorize existing GenAI applications employed, their practical roles in curriculum development within the tertiary health education setting and the perceptions of stakeholders (students and educators) towards these applications. By synthesizing themes from existing literature on the recent GenAI applications in tertiary health education, this review aimed to identify key trends and concerns in the literature that can guide future directions on how to integrate GenAI applications smoothly and effectively into the tertiary health education sector.

Methods

Study Design

This study is a scoping review conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews checklist (PRISMA ScR) (Peters et al., 2020; Tricco et al., 2018). The protocol of the scoping review was guided by the five-stage framework for scoping reviews (Arksey & O'Malley, 2005). The Population, Concept, Context (PCC) framework was utilized to help construct clear eligibility criteria and objectives of the project (Pollock et al., 2023). The protocol of this project was registered with the Open Science Framework (Cheung et al., 2025).

A scoping review was selected to address the research question since it was useful in answering broad questions, identifying gaps in literature and summarizing findings from a broad area of knowledge (Pollock et al., 2023). By mapping information from all types of study designs as well as grey literature, a scoping review was believed to effectively break down the broad and complex topic of the integration of GenAI into health education. Its comprehensive summary of the available literature could also provide breadth into the chosen research topic and help guide future research. We identified the following research question to guide the search: How is GenAI being integrated into tertiary health education for health professional students?

Eligibility Criteria

The inclusion and exclusion criteria were established using the PCC (Population, Concept, Context) framework:

- Population: Tertiary level students and educators in health and medical disciplines
- Concept: The integration of GenAI into tertiary health education
- Context: Tertiary health education sector (undergraduate level or above)

The population of interest was the tertiary level students and educators in health, and medical disciplines. These disciplines include doctors, nurses, and allied health practitioners. As defined by Allied Health Professions Australia, allied health practitioners include art therapists, audiologists, nutritionists, chiropractors, counsellors/psychotherapists, diabetes educators, dieticians, exercise physiologists, genetic counsellors, medical radiation practitioners, music therapists, occupational therapists, optometrists, orthoptists, prosthetists, osteopaths, paramedics, pedorthists, perfusionists, pharmacists, physiotherapists, podiatrists, psychologists, rehabilitation counsellors, social workers, sonographers, and speech pathologists (Allied health professions

Australia, n.d.). Studies involving any of the above disciplines were eligible for inclusion. Various inclusion and exclusion criteria were put in place, as demonstrated in Table 1.

Table 1

Table of Inclusion and Exclusion Criteria.

Inclusion Criteria	Exclusion Criteria
Peer-reviewed studies Empirical studies with all types of study designs Studies in English Grey literature	Editorials, review articles, websites, letters to editor Papers with only abstract (no full text) Studies not in English
Tertiary health education Medical education	Studies not focusing on the tertiary education sector Studies not focusing on GenAI, but only artificial intelligence

Types of Evidence Sources

This review included empirical research articles with various types of study designs, including but not limited to randomized controlled trials, experimental, qualitative and descriptive studies, if they focused specifically on GenAI and the tertiary level of health education. Conceptual, theoretical, commentary, reviews and opinion papers were excluded unless they presented original data relevant to GenAI use in tertiary health education. Editorials, websites, opinion pieces, conference abstracts, and non-English studies were excluded. Literature search was limited to publications from 2021 onwards to maintain currency of sources.

Since the research question was relatively broad, the scoping review included original research of all types of study designs and peer-reviewed sources to maximize data yield and the mapping of information. Review articles, editorials, websites and letters to editor were excluded. A preliminary attempt at searching for articles in languages other than English was made, yielding three in Japanese, one each in Chinese, French and German. However, free translation tools such as Google Translate and DeepL Translator proved unreliable in translating the true meaning of statements. Since the project did not have funding for a translator, it was decided that articles not in English were excluded. Furthermore, sources that addressed artificial intelligence instead of GenAI were excluded as they were not specific to the identified concept. Similarly, sources unrelated to the tertiary or higher education sector were excluded due to a lack of specificity towards the context.

The release of ChatGPT in November 2022 was the major catalyst for the growing number of GenAI tools developed and applied in the tertiary health education sector. However, the research team decided to begin the literature search from 2021 to ensure that relevant publications prior to the introduction of ChatGPT were also captured.

Search Strategy

A three-step search strategy was used to develop the search terms with the help of an academic librarian. Initial search on Ovid Medline helped to develop the final search terms. The final search terms, including Boolean operators, were developed (See Table 2). This comprised keywords and subject headings related to GenAI and tertiary health, and medical education such as "Artificial Intelligence" OR "Generative AI" OR "Generative Artificial Intelligence" OR "Machine learning" OR chatgpt OR chatbot OR "Large Language Model", AND "Health Education" OR "Medical Education" OR "Healthcare Student" OR "Medical Student" OR "Allied Health Student" OR "Allied Health Education", AND "Tertiary Education" OR undergraduate OR postgraduate OR "Higher education" OR “university”. Finally, five databases, Ovid Medline, Ovid Emcare, Scopus, CINAHL, and ERIC, were systematically searched for English articles published from 2021 to October 2024. The database search was conducted by the first author on 15 November 2024. The complete search string for all the databases is available as an open access dataset (Cheung et al., 2025).

Table 2

Final Search Terms

PCC	Concepts	Keywords	Subject Headings
Population	Tertiary education students and educators in health disciplines	Health educa*, Medical educa*, allied health educat*, healthcare student*, medical student* and allied health student*	Students, Health Occupations Students, Medical Health Educators Allied Health- Personnel Allied Health- Occupations
Concept	Integration of GenAI (into tertiary health education)	Generative artificial intelligence, generative AI, generative language model, artificial intelligence, machine learning, virtual, simulat*, ChatGPT, chatbot, automatic scoring, AI literacy and integrat*	Artificial Intelligence Machine Learning Simulation training
Context	Tertiary health education	tertiary health educat*, undergrad* and postgrad*	Health Education Education, Medical

Screening

The database search results were added to the Covidence systematic review software (2024) for removing duplicates and screening. A total of 598 studies underwent title and abstract screening, followed by full text screening. Two reviewers independently assessed records against the inclusion and exclusion criteria. A third author resolved the discrepancies that arose in abstract and full text screening. All conflicts were resolved through discussion and reaching a consensus among the research team. The inclusion and exclusion criteria were further refined before full-text review proceeded. The inclusion criteria stated that articles must focus on GenAI and the tertiary health education sector. Only those articles which met the inclusion and exclusion criteria with a 100% agreement rate between reviewers were included in the final tally.

Second Iteration of the Search Strategy

A second search was conducted on July 19, 2026 on SCOPUS by the corresponding author. The search string for SCOPUS was used from the search string used in the initial search after editing the dates to search for articles from November 2024 to July 2026. Additionally, a manual search of the first 50 articles on Google Scholar was also conducted by the first author. The second search was targeted rather than replicating the full database search. The search strategy, databases, date, and eligibility criteria are reported as per guidelines (Peters et al., 2020; Tricco et al., 2018) to ensure transparency and reproducibility. This consensus decision by the project team reflected the rapidly evolving AI literature. The supplementary search was intended to identify additional relevant literature rather than to claim exhaustive retrieval of all publications within this continually developing field. The search strings for Google Scholar and SCOPUS are given in Appendix A. Articles were screened for title and abstract by two reviewers as per inclusion and exclusion criteria. The articles for full-text screening were distributed to all authors such that at least two reviewers screened the articles independently. Any discrepancies were resolved by the third reviewer. Only those articles which met the inclusion and exclusion criteria with a 100% agreement rate between reviewers were included in the final tally.

Data Extraction

Data from included articles was extracted and entered in a Microsoft Excel spreadsheet in a study characteristics table to answer the research question, which included, the study, country, discipline, study design, application of GenAI, practical application and key findings. Data extraction was initially piloted for 3 articles between the first author and the corresponding author. After reaching more than 95% agreement on the data, the remaining data extraction was conducted by the first author and cross-checked by all the remaining authors. Disagreements were settled by discussion and consensus with the research team. The data extraction was consistent with PRISMA guidelines for scoping reviews; thus, the risk of bias was not assessed (Tricco et al., 2018).

Data Analysis

Subheadings included the following aspects: the characteristics of studies (publication year, country, sample size, course discipline), category of application, study design, and key findings of the studies. Categories were identified to guide the reporting themes, which were finalized after consensus by all authors.

Results

The search and selection process are reported using a PRISMA-ScR (Peters et al., 2020; Tricco et al., 2018) flow diagram (See figure 1). The initial search yielded 1112 titles, of which 514 were duplicates. After removal of duplicates, 598 studies were available for screening. Following title and abstract screening, 524 articles were removed as they did not specifically focus on GenAI, but only general AI. A total of 74 studies were included for full text screening. Full text screening resulted in removing studies which focused on the test performance of GenAI tools, instead of the integration of such tools into tertiary health education. These articles were removed after consensus among authors. A total of 13 articles were included in the final analysis. The second search, conducted in July 2026, yielded 402 articles for title and abstract screening. This resulted in 23 articles for full text screening. Finally, seven articles were added, bringing the

final tally to 20. The extracted data was categorized to answer the research question: “How is GenAI being integrated into tertiary health education?”.

Characteristics

As summarised in Table 3, the included articles were published between 2023 and 2026, reflecting the recency of the integration of GenAI into tertiary health education. Geographically, the 20 studies spanned across North America, Asia, Europe, and the Middle East. This includes the USA (n=5), China (n=5) and one each from Thailand, Taiwan, Germany, Switzerland, Bahrain, Israel, India, Turkey, Saudi Arabia and Hong Kong Special Administrative Region (HKSAR)/Singapore/Ireland/UK. 80% of studies focused on medical education (n=16), with the remainder covering nursing, anatomy, dentistry and healthcare management education. Most studies utilized ChatGPT (n=16), while two applied platform-specific adaptations, namely AnatomyGPT (Collins et al., 2024) and Glass AI (Robleto et al., 2024). Various methodologies were employed, which included randomised controlled trials (n=3), descriptive (n=5), prospective (n=3), experimental (n=1), quasi-experimental (n=6), and mixed-methods (n=2) designs. The reported methodologies of the included studies are intended to be mutually exclusive. The comprehensive data extraction table is available as an open access dataset for the initial search conducted in 2024 (Cheung et al., 2025).

Figure 1

PRISMA ScR Flowchart (Peters et al., 2020; Tricco et al., 2018)

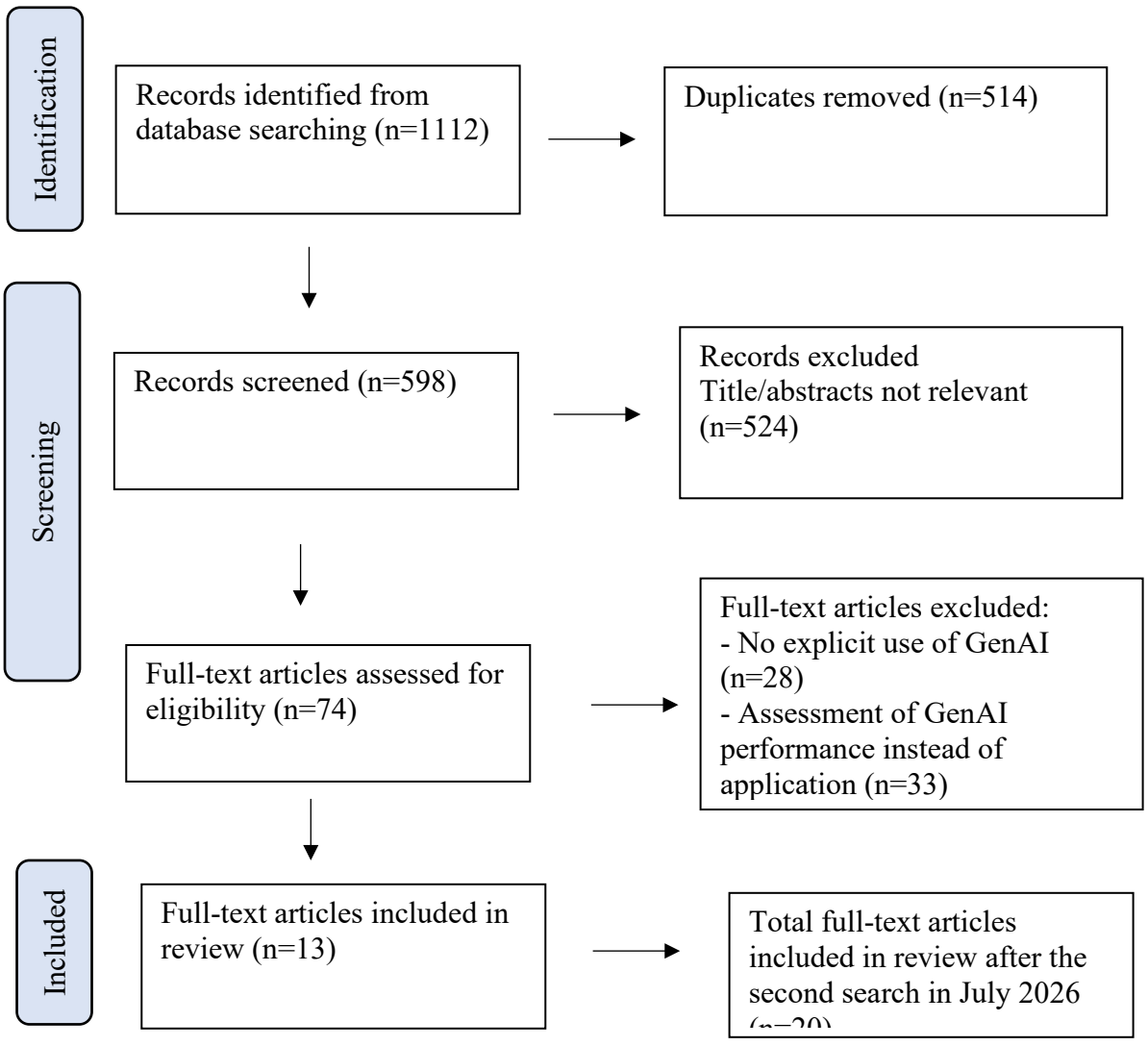


Table 3

Characteristics and key findings of all the included studies from the initial and the second search

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
1. Burke et al. (2024)	USA	Medical education	Cross-sectional descriptive	Scoring free-text clinical notes of medical students by ChatGPT	Automated grading, assessment generation, and feedback.	ChatGPT achieved a markedly lower incorrect scoring rate (1.0%) than the standardized rate (7.2%), suggesting LLMs can reliably score student notes in real time while providing immediate performance feedback.	
2. Chang et al. (2024)	Taiwan	Nursing education	Post-test quasi-experimental	Incorporation of CIDI (Clarify, Ideate, Develop, Implement) model-based ChatGPT system to promote critical thinking and problem-solving skills of students.	Improved learning outcomes	Learners using the CIDI model–based ChatGPT system (mean = 4.53; SD = 0.71) showed significantly greater gains in problem-solving skills compared with conventional methods (mean = 3.75; SD = 0.58). The interactive design promoted critical thinking by enabling conversational practice and judgment-based decision-making.	
3. Cheung et al. (2023)	HKSAR, Singapore, Ireland, and UK	Medical education	Comparative quasi-experimental	ChatGPT-generated multiple choice questions (MCQs) for medical examinations	Automated grading, assessment generation, and feedback	The time spent by ChatGPT to create 50 MCQs was 20 minutes 25 seconds, compared to 211 minutes 33 seconds by human examiners.	Reliability issues and potential bias in training data

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
4. Chiu et al.	USA	Medicine	Quasi-experimental	Breaking bad news training using predesigned scenarios in ChatGPT	Teach medical trainees the skills for breaking bad news.	Utilizing ChatGPT as a tool for communication skills training can enhance medical students' confidence levels and increase their perceived communication skills and professionalism particularly to patients	Privacy concerns
5. Cicek et al, 2024	Turkey	Medicine	Randomized controlled trial	To evaluate the effectiveness of ChatGPT-generated feedback compared to expert-written feedback in improving clinical reasoning skills among first-year medical students.	ChatGPT as supplementary tool in resource-limited settings where expert feedback is not readily available.	There was no significant difference between the mean scores of the control and experiment groups in overall performance on Key-Features Questions immediately or after 10 days. However, the control group outperformed the ChatGPT group in complicated urinary tract infection cases ($P < .001$).	While the effect of ChatGPT feedback was generally on par with the effect of expert feedback, the study identified limitations in AI-generated explanations for more nuanced diagnosis and treatment.
6. Collins et al. (2024)	USA	Anatomical Sciences education	Exploratory descriptive	Customized GenAI platform: AnatomyGPT, for anatomy education	Improvement in learning outcomes	AnatomyGPT provided more source citations than standard GPT when generating responses.	Confidentiality of materials uploaded to AnatomyGPT
7. Daniel et al (2026)	India	Medicine	Longitudinal, prospective non-randomized delayed intervention	to evaluate supervised ChatGPT-assisted active recall of previously learned material and its	As a cognitive scaffold to provide feedback & refine reasoning in real time. Improved self-awareness by prompting metacognitive	Both lower and high-achieving students benefited from the intervention, but the effect was more pronounced in the lower-achieving-learner group. Student feedback reflected high satisfaction, strong cognitive	Over-reliance, fears about reduced independent thinking or exposure to AI-generated inaccuracies.

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
				durability over time	reflection and boosting confidence.	engagement, and increased motivation.	Challenges with the clarity and flow of AI–student interactions.
8. Elsamanoudy (2026)	Saudi Arabia	Medicine	quasi-experimental mixed-methods educational evaluation (non-randomized, pragmatic intervention	to assess the effectiveness and usability of a customized AI-powered tutoring system, GPT Reproductive Physiology Companion, among third-year medical students enrolled in the Reproduction module.	As a safe and supportive supplement to traditional medical education, rather than replacing established instructional practices.	High levels of student satisfaction across various domains, particularly in conceptual understanding, formative assessment, and critical thinking support. Students consistently reported that the tool was user-friendly, aligned with curricular content, and contributed positively to their learning experience, regardless of gender, academic background, or GPA.	
9. Gao et al. (2026)	China	Medicine	Prospective, randomized controlled trial	As a conversational simulated patient using an AI-based web-based interface	Improve students' history-taking performance, particularly in completeness and organizational efficiency gathering.	The AI group achieved higher checklist completion, missed fewer key items, and obtained higher total OSCE scores than controls, with large effect sizes. They also performed better in medical interviewing and organization/efficiency, No significant differences were found in professionalism, clinical judgment, counselling skills, or overall clinical competence.	Privacy concerns

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
10. Holderried et al. (2024)	Germany	Medical education	Prospective observational	ChatGPT-simulated patient history-taking with automated feedback on history-taking skills of medical students	Automated grading, assessment generation, and feedback	The role-play and responses simulated by GPT-4 chatbot was medically plausible in more than 99% of cases, demonstrating effectiveness in providing comprehensive feedback on history-taking dialogs.	Unexpected behaviors, escape prompts, and increasing reliance on GenAI by students were raised.
11. Huang et al. (2025)	China	Dentistry	Randomized Controlled Trial	Video learning with ChatGPT-3.5 for skill acquisition	Dental skill learning	The ChatGPT group showed better performance ($p<0.01$), greater self-efficacy ($p=.04$) and motivation ($p=.02$), and lower cognitive load ($p<0.01$) than controls. Benefits were greatest among learners with low spatial ability.	
12. Laye & Wells (2024)	USA	Medical education	Practical descriptive	Formation of balanced student groups based on academic background and demographics using ChatGPT analysis	Improvement in learning outcomes	ChatGPT created balanced groups with greater synchronicity in completing class assignments, broader group participation, all under shorter group construction time. The application addressed past issues of time-consuming group formation by human staff and the imbalance of students in terms of prior experience and medical knowledge within groups.	

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
13. Li et al. (2024)	China	Medical education	Mixed methods	Application of ChatGPT to enhance the writing of English papers by non-native medical students through automatic feedback	Automated grading, assessment generation, and feedback Student perceptions	ChatGPT-assisted instruction improved students' coherence, grammar and vocabulary, increasing scores by 4.23 points and receiving 92% positive feedback for translation and real-time correction. This suggests strong potential for supporting assignment grading and reducing workload.	Authenticity and credibility of information from ChatGPT.
14. Naamati-Schneider (2024)	Israel	Healthcare management education	Explanatory descriptive	Integration of ChatGPT to problem-based learning framework to build AI and digital literacy of students	Improvement in learning outcomes Student perceptions	Students found ChatGPT highly accessible, user-friendly, supporting autonomous learning, and time efficient. Some students perceived ChatGPT as an opportunity to improve English language skills.	Inaccuracies, difficulty in phrasing clear and specific prompts, importance of critically evaluating and cross-referencing responses with other academic sources
15. Prachnakorn et al. (2024)	Thailand	Medical education	Experimental	Application of ChatGPT to transform traditional lectures to interactive workshops to enhance scientific report-writing skills of preclinical	Improvement in learning outcomes	ChatGPT successfully assisted authors in transforming a didactic lecture to a workshop, helping instructors to summarize students' inputs and increased engagement in the workshop.	Questionable reliability of GenAI, a hybrid approach combining human intellect and AI competence recommended.

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
				medical students			
16. Robleto et al. (2024)	USA	Medical education	Prospective	Application of 'Glass AI', a platform that combined GPT-4 with clinical guidelines	Improvement in learning outcomes Student perceptions	96% of participating medical students stated Glass AI increased their confidence in making the diagnosis. 97% stated Glass AI was easy to employ and effective in helping them reach the correct diagnosis.	Patient privacy, legal liability, and biases in the GenAI algorithms, including age, socioeconomic status, gender, and race.
17. Song et al. (2025)	China	Medicine	Quasi-experimental	Integration of GenAI for interactive teaching strategies in an undergraduate nursing course	Improvement in learning outcomes	The experimental group reported significantly better higher-order thinking scores (P= 0.018) and AI literacy (P<0.001) than the control group after the course	
18. Sridharan et al. (2024)	Bahrain	Medical education (pharmacology)	Descriptive	Integration of GenAI to enhance classroom instructions and student assessment	Improvement in learning outcomes	GenAI tools were useful in helping plan instructional methods in medical education programs, generating useful specific learning outcomes (SLOs) and detailed instructions.	Extracting details from unauthenticated sources. Test items created from GenAI should be reviewed by faculty members for content validity.
19. Thomae et al. (2024)	Switzerland	Medical education	Mixed methods	Integrating ChatGPT into a blended learning course, with student evaluation.	Improvement in learning outcomes and student perceptions	The integration yielded high satisfaction scores. Students found extended prompts generated information of higher quality, flagging the importance of students' ability to write specific and refined prompts.	Data protection issues

Study	Country	Discipline	Study design	Application of GenAI	Practical application	Key findings	Concerns
20. Wu et al. (2024)	China	Medical education	Controlled quasi-experimental	Integration of ChatGPT into clinical teaching of hepatobiliary surgery.	Improvement in learning outcomes	Improved students' understanding of theoretical knowledge and clinical skills. The experimental group obtained significantly higher final theoretical assessment scores than the control group	

Categories of Findings

Across the studies, GenAI was predominantly used to support formative assessment through automated feedback, evaluation of clinical reasoning, as a companion chatbot, and assessment of communication and procedural skills. As derived from the extracted data in Table 3, the practical applications of GenAI will be presented in three broad categories:

1. Applications of GenAI in Tertiary Health Education
2. Perceptions of GenAI use in health education
3. Ethical and Responsible Use of GenAI

Applications of GenAI in Tertiary Health Education

Application in Assessment

In terms of the assessment-related GenAI applications employed, GenAI was applied in seven studies for automated grading, assessment generation, and feedback (Burke et al., 2024; Cheung et al., 2023; Çiçek et al., 2024; Daniel et al., 2026; Elsamanoudy et al., 2026; Holderried et al., 2024; Li et al., 2024). Two studies (Burke et al., 2024; Holderried et al., 2024) employed ChatGPT to assess free-text clinical notes or simulate patient scenarios with immediate feedback provided; both studies reported high accuracy and reliability of the GenAI tool, with a key benefit being real-time detailed feedback generated by GenAI to students. One study utilized ChatGPT to automatically score and assist non-native medical students in writing English papers, where immediate feedback from ChatGPT led to improved student scores and writing quality (Li et al., 2024). The study by Cheung et al. (2023) utilized ChatGPT to generate MCQs for medical examinations, finding the tool to be highly time-efficient and capable of generating MCQs of comparable quality to human examiners. Two studies highlighted that automated grading reduced staff workload, freeing up time for more personalized engagement with students (Cheung et al., 2023; Li et al., 2024). Çiçek et al. (2024) demonstrated that ChatGPT-generated feedback improved students' clinical reasoning performance, although expert feedback remained superior in providing nuanced clinical judgement. GenAI also enabled immediate, iterative feedback that promoted self-reflection and independent learning, as reported by Daniel et al. (2026), when integrating ChatGPT to evaluate training in retrieval skills. Overall, these studies suggest that through automated grading and feedback, GenAI can enhance educators' productivity and student assessment experiences.

Application in Instruction and Teaching

The primary educational application across thirteen studies was GenAI functioning as an interactive tutor or learning companion. ChatGPT supported communication skills training (Chiu et al., 2025), knowledge retrieval tutorials (Daniel et al., 2026), physiology education through a customized AI tutor (Elsamanoudy et al., 2026), simulated patient interactions for gynecological history taking (Gao et al., 2026), clinical skills instruction in dental education (Huang et al., 2025), clinical reasoning (Çiçek et al., 2024), and development of higher-order thinking in nursing education (Song et al., 2025). These conversational AI systems provided personalised explanations, immediate feedback, unlimited opportunities for practice, and flexible access beyond classroom hours. Li et al. (2024) in their study reported improved students' coherence, grammar and vocabulary through ChatGPT-assisted instruction for English as second language

students. Students valued the ability to engage repeatedly with realistic clinical scenarios in a psychologically safe environment without fear of judgement.

Among seven of the included articles, there was a consistent trend of using GenAI tools to enhance engagement, collaboration, and critical thinking of students in tertiary health education courses (Chang et al., 2024; Collins et al., 2024; Laye & Wells, 2024; Prachnakorn et al., 2024; Sridharan et al., 2024; Thomae et al., 2024; Wu et al., 2024). Two studies described the transformation of traditional classroom instructions into more engaging, student-centered formats, with GenAI supporting brainstorming and planning for classes (Prachnakorn et al., 2024; Sridharan et al., 2024). The study by Laye & Wells, (2024) explored the use of GenAI to form balanced student groups, reporting shorter group construction time and broader group participation. Another study utilized ChatGPT to develop a customized GenAI platform, which provided anatomy-specific education responses to students (Collins et al., 2024). Overall, these studies suggest that integrating GenAI into course designs can lead to enhancement of learning experiences and student improvements.

Application in Course Design

Several studies demonstrated that GenAI could support curriculum development through incorporating theoretical frameworks, creation of authentic learning activities, simulated patients, structured tutorials, and adaptive learning resources. Gao et al. (2026) illustrated how GenAI can efficiently create simulated patients for history-taking exercises, while Elsamanoudy et al. (2026) embedded a customized ChatGPT physiology companion into a reproductive medicine module. Daniel et al. (2026) integrated AI-assisted knowledge retrieval sessions into existing curricula. In nursing education, Song et al. (2025) incorporated GenAI into learning activities that assessed higher-order thinking, demonstrating improvements in students' analytical and problem-solving abilities alongside AI literacy. Five studies integrated GenAI into teaching units and learning designs (Chang et al., 2024; Naamati-Schneider 2024; Robleto et al., 2024; Thomae et al., 2024; Wu et al., 2024) and reported significantly higher student learning improvement in terms of critical thinking and clinical skills. These findings indicate that GenAI can enhance curriculum scalability and reduce faculty workload when educational resources are carefully designed and reviewed by educators.

Perceptions of GenAI Use in Health Education

Perceived Benefits

A consistent finding across all studies was the positive perception of GenAI as a learning support tool. Students reported that AI enhanced engagement, motivation, confidence, and learning satisfaction. Personalised feedback and real-time responses allowed learners to progress at their own pace and revisit concepts as needed (Daniel et al., 2026; Elsamanoudy et al., 2026). GenAI also promoted active learning by encouraging questioning, reflection, and self-directed knowledge construction. Studies involving simulated patients (Gao et al., 2026) and communication skills (Chiu et al., 2025) demonstrated that AI provided safe environments for repeated practice of interpersonal skills before clinical placement. In nursing education, Song et al. (2025) further showed that AI use improved higher-order thinking skills and AI literacy, preparing students for technology-enhanced healthcare environments. Faculty similarly recognized AI’s potential to increase efficiency, expand learning opportunities, and provide scalable educational support. Four studies focused largely on student perceptions (Li et al., 2024; Naamati-Schneider, 2024; Robleto et al., 2024; Thomae et al., 2024). In all four studies, students

found ChatGPT and other GenAI tools highly accessible and user-friendly, supporting autonomous learning (Li et al., 2024; Naamati-Schneider, 2024; Robleto et al., 2024; Thomae et al., 2024). Li et al. (2024) reported 92% positive student feedback on ChatGPT as a translation and grammar correction tool; while 96% of participating medical students in the study by Robleto et al. (2024) found that GenAI increased their confidence in making diagnoses.

Perceived Risks and Barriers

Despite widespread optimism, all studies identified important limitations. The most reported concern was the potential for inaccurate or fabricated information (hallucinations), requiring students to critically evaluate AI-generated responses rather than accepting them uncritically. The study by Naamati-Schneider (2024) reported that students were sometimes surprised by the inaccuracies of ChatGPT, indirectly reminding them of the importance of critically evaluating and cross-referencing GenAI information with other academic sources. Thomae et al. (2024) added that student satisfaction depended on their ability to write specific and well-structured prompts. Overall, the studies (Li et al., 2024; Naamati-Schneider, 2024; Robleto et al., 2024; Thomae et al., 2024) suggest that students are mostly open to integrating GenAI into their learning but require foundational training in prompt literacy and appraisal skills to utilize GenAI tools effectively. Participants also expressed concerns regarding overreliance on AI, reduced independent critical thinking, and the possibility that students might substitute AI assistance for genuine learning (Daniel et al., 2026). Technical challenges, inconsistent response quality, limited contextual understanding, and occasional superficial explanations were additional barriers reported across the studies. All studies generally concluded that AI should augment rather than replace educator expertise.

Ethical and Responsible Use of AI

Despite the overall positive outlook, concerns of potential drawbacks were reported in 80% (n=16) studies, spanning across ethical, educational, and practical domains. Privacy considerations were particularly relevant for studies involving simulated clinical interactions. Gao et al. (2026) and Chiu et al. (2025) highlighted the importance of avoiding identifiable patient information when interacting with publicly available AI systems. Institutional governance, secure AI platforms, and adherence to ethical standards were considered necessary to protect confidentiality during educational activities. Three further studies raised concerns about data privacy and confidentiality, particularly when uploading sensitive academic or patient information (Collins et al., 2024; Robleto et al., 2024; Thomae et al., 2024).

Concerns about bias in training data and misinformation were flagged in six studies (Cheung et al., 2023; Holderried et al., 2024; Li et al., 2024; Prachnakorn et al., 2024; Robleto et al., 2024; Sridharan et al., 2024). Two studies highlighted potential student overreliance on GenAI for answers, reducing independent reasoning and content verification (Holderried et al., 2024; Li et al., 2024). In addition, three studies reported that students often struggled to write effective prompts, indicating the need for better instructions in GenAI literacy and structuring prompts (Li et al., 2024; Naamati-Schneider, 2024; Thomae et al., 2024).

Across all studies, accountability remained firmly with educators rather than AI systems. Faculty retained responsibility for ensuring the accuracy of educational materials, maintaining assessment integrity, monitoring student learning, and validating AI-generated feedback. Rather than replacing teachers, AI was consistently positioned as a supportive educational partner

whose outputs require human oversight, professional judgement, and continuous quality assurance.

Discussion

This scoping review demonstrates rapidly growing evidence that GenAI can enhance teaching, learning, and assessment across health professions education. Across the twenty included studies, GenAI, specifically ChatGPT, was most implemented as an intelligent tutor, simulated patient, personalised learning assistant, and feedback provider. These applications supported communication skills, clinical reasoning, knowledge acquisition, procedural training, and self-directed learning. Students generally perceived GenAI as engaging and accessible, while educators recognized its potential to improve educational efficiency, personalised feedback, and curriculum scalability. Despite these benefits, the studies consistently emphasized that GenAI should complement rather than replace educators, with human oversight remaining essential to ensure educational quality, clinical accuracy, and responsible implementation.

These findings align with previous reviews that advocate institution-wide strategies for GenAI adoption, including faculty development, AI literacy, governance frameworks, and clear academic integrity policies (Koontz, 2025; Cordero et al., 2024). Collectively, the evidence suggests that the educational value of GenAI depends less on the technology itself than on intentional instructional design. Rather than embedding AI as an isolated innovation, educators should integrate it within structured learning activities that promote reflection, critical thinking, authentic clinical reasoning, and real-world application (Crompton et al., 2026).

An important finding is that the evidence base is in its infancy. Although the search strategy included studies from 2021 onwards, all eligible publications were published after 2023, reflecting the rapid emergence of ChatGPT and related large language models. Consequently, much of the current literature focuses on feasibility, learner perceptions, and early implementation rather than rigorous evaluation of learning outcomes, clinical competence, behavioral change, or patient outcomes. There remains a need for longitudinal, experimental, and multi-institutional research to determine the sustained educational effectiveness of GenAI and identify best practices for curriculum integration (Fui-Hoon Nah et al., 2023).

Another notable finding is the dominance of medical education within the literature, with comparatively limited research conducted in allied health disciplines. This imbalance likely reflects the highly regulated nature of allied health education, where curriculum innovation must align with accreditation standards, competency frameworks, and patient safety requirements. As AI technologies become increasingly integrated into healthcare delivery, expanding research across allied health professions represents an important priority to ensure graduates develop AI capabilities appropriate to their clinical contexts (Crompton et al., 2026).

The above gaps are consistent with the review study on cross sector analysis by Eddine et al. (2025). They reported a disparity in adoption of ChatGPT across various disciplines, with high usage reported across Computer science and social sciences and relatively lower in health disciplines (Eddine et al., 2025). Of the health disciplines, medical education dominated in their results as well (Eddine et al., 2025). Atif et al. (2026), in their recent study, provide the first systematic, national-level analysis of Australian university alignment with TEQSA's guiding

principles for GenAI in assessment, offering a baseline against which future policy developments can be measured. They evaluated discipline-specific focus on information and communications technology education and identified a gap that GenAI guidance for technical assessments, is largely absent across Australian universities (Atif et al., 2026).

One of the most significant contributions of this review is the identification of AI literacy as an emerging core competency. While earlier literature has largely focused on the benefits and risks of GenAI, relatively little attention has been given to preparing students and educators to use these technologies critically and responsibly. The findings suggest that AI literacy extends beyond technical proficiency to include understanding AI limitations, recognizing bias and hallucinations, evaluating the reliability of AI-generated information, and applying ethical judgement within clinical decision-making. Frameworks such as the CSSD AI Literacy Framework and recent recommendations for assessment redesign (Al-Shawee et al., 2026; Atif et al., 2026; Cordero et al., 2024; Koontz, 2025) reinforce the need for AI literacy to become embedded throughout health professions curricula rather than taught as a standalone topic.

The review also reinforces ongoing concerns regarding academic integrity, bias, transparency, privacy, and overreliance on AI-generated content. Several studies highlighted the potential for AI systems to perpetuate biases present in training data, influencing educational resources, simulated patients, and clinical recommendations. These findings underscore the importance of institutional governance, transparent AI use, educator validation of AI-generated outputs, and explicit teaching that AI outputs require critical appraisal rather than unquestioning acceptance.

Implications

The findings have several important implications for health professions education. As AI becomes increasingly integrated into clinical practice, graduates must develop the knowledge, skills, and judgement required to work effectively alongside AI technologies. Consequently, curricula should move beyond teaching students how to use AI tools and instead focus on developing AI literacy, critical appraisal, ethical reasoning, and responsible clinical decision-making.

Successful implementation requires a whole-of-institution approach that includes curriculum redesign, faculty development, governance frameworks, and assessment reform. AI should be embedded within authentic learning experiences where students critically evaluate AI-generated outputs, compare them with evidence-based practice, and reflect on their strengths and limitations. Faculty development is equally important, enabling educators to design AI-enabled learning activities, develop AI-resilient assessments, and model responsible AI use.

For assessment, the findings support a hybrid approach in which GenAI is primarily used to enhance formative learning through personalised feedback, simulated clinical interactions, and deliberate practice, while educators retain responsibility for summative assessment, quality assurance, and professional judgement. Future assessment design should increasingly evaluate students' ability to critique, verify, and appropriately integrate AI-generated information rather than simply reproduce knowledge.

From a research perspective, future studies should prioritize robust evaluations of long-term educational outcomes, graduate preparedness, clinical competence, and patient-related outcomes. Greater attention is also needed to allied health disciplines, where evidence remains limited despite the growing integration of AI into clinical practice. Developing and evaluating accreditation-informed models for AI integration will be essential to ensure innovation occurs alongside patient safety, professional standards, and public trust.

Overall, this review suggests that the future of GenAI in health professions education lies not in replacing educators, but in augmenting learning through carefully designed, ethically governed, and pedagogically informed integration that prepares graduates for AI-enabled healthcare systems.

Limitations

As with any scoping review, this study has certain limitations that should be considered. Firstly, there is a risk of selection bias as the search strategy is limited to English literature only. It is possible that the database searches did not obtain all relevant literature as set by the eligibility criteria. Secondly, the search was initially limited to the period between 2021 and October 2024, meaning that some literature regarding the development of GenAI prior to the release of ChatGPT might have been excluded. This was mitigated to some extent by a second search on limited databases in July 2026. However, due to the limiting nature of the search, relevant current articles in other databases may have been omitted, while the most recent developments of GenAI, particularly the emergence of the DeepSeek GenAI model, were not captured. The rapidly evolving nature of GenAI technologies means that the findings may not have fully reflected the capabilities of the latest GenAI tools. Thirdly, the data extraction may have involved a degree of interpretive variability. Although the themes were developed through review and consensus among the research team, decisions regarding how findings were grouped may have been influenced by researcher judgement. As such, alternative categories may have been possible, particularly given the overlap between GenAI applications, perceived benefits, risks, and ethical considerations.

Collectively, the studies in this review suggest that GenAI is most valuable as a complementary educational technology that enhances personalised learning, formative assessment, and curriculum development. Its successful integration into health professions education requires robust governance, educator oversight, AI literacy training, and ethical safeguards to maximize educational benefits while mitigating risks associated with accuracy, bias, privacy, and inappropriate dependence on AI. Rather than drawing definitive conclusions about the effectiveness of GenAI in tertiary health education, this review illuminated where research currently stands and where it should go. Most studies were exploratory and focused on short-term implementations, leaving questions about long-term effects of GenAI on student learning and the curriculum. As this field is still in an experimental phase, this review highlighted the importance of directing future research into long-term evaluations, particularly in terms of designing AI literacy training modules and ethical guidelines.

Conclusion

This scoping review mapped the current literature on the integration of specific GenAI applications into tertiary health education. The findings demonstrated that GenAI is being rapidly integrated into tertiary health education as a supportive tool for automated grading and feedback, as well as enhanced learning outcomes and improvements in student engagement and critical thinking. Although the majority of GenAI applications employed were well-received, concerns of reliability, ethical use, and over-reliance were often cited. A key knowledge gap was identified in the lack of AI literacy training within current tertiary health education programs, which could risk underpreparing graduates. As the influence of GenAI continues to grow in the tertiary health education sector, it is imperative for educational institutions to develop AI literacy training workshops and ethical guidelines. AI literacy training modules should feature prompt literacy, informed skepticism on the quality of GenAI responses and awareness of academic integrity. Furthermore, longitudinal research on the long-term impacts of GenAI and the specific design of AI literacy training workshops is recommended.

Declarations

Ethics Statement

Due to the nature of this review, ethical approval and informed consent were not required. The project did not have any major risks as it did not involve human research and only used literature that had already been published (National Health and Medical Research Council, 2018).

Funding Statement

This research received no external funding.

AI Use Statement

Generative artificial intelligence was used for grammar and spelling using the free version of Grammarly. Generative artificial intelligence was not used to generate, analyze, or interpret research data, nor to create original scholarly content. All content was reviewed, verified, and approved by the authors, who take full responsibility for the accuracy and integrity of the work.

Conflicts of Interest Statement

The authors have no conflicts of interest related to the research, authorship, or publication of this article.

References

Alam, F., Lim, M. A., & Zulkipli, I. N. (2023). Integrating AI in medical education: Embracing ethical usage and critical understanding. *Frontiers in Medicine*, 10. https://doi.org/10.3389/fmed.2023.1279707

Allied Health Professions Australia. (n.d., first accessed May 9, 2023). *Allied health professions*. https://ahpa.com.au/allied-health-professions/

Al-Shawee, N., McElvaney, G., Strawbridge, J., & Spooner, M. (2026). When assessment theory meets generative AI: Reimagining SBA design in medical education. Perspectives on *Medical Education*, 15(1), 239–250. https://doi.org/10.5334/pme.2033

Atif, A., Jha, M., Richards, D., & Cibin, A. (2026). Institutional alignment with TEQSA's generative AI principles in Australian universities: A national document analysis with implications for ICT education. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.11001

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, G. M. (2025). Implementing generative AI (GenAI) in Higher education: A systematic review of Case studies. *Computers and Education: Artificial Intelligence*, 8, 100407. https://doi.org/10.1016/j.caeai.2025.100407

Benmamoun, M. (2023). Reinventing international business education: integrating the power of generative AI. *AIB Insights*. https://doi.org/10.46697/001c.90397

Boscardin, C., Gin, B., Golde, P. B., & Hauer, K. E. (2023). ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. *Academic Medicine*, 99(1). https://doi.org/10.1097/acm.0000000000005439

Burke, H. B., Hoang, A., Lopreiato, J. O., King, H., Hemmer, P., Montgomery, M., & Gagarin, V. (2025). Assessing the ability of a large language model to score free-text medical student clinical notes: quantitative study. *JMIR Medical Education*, 10, e56342. https://doi.org/10.2196/56342

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *Internation Journal of Educational Technology in Higher Education*, 20(1). https://doi.org/10.1186/s41239-023-00411-8

Chang, C.-Y., Yang, C.-L., Jen, H.-J., Ogata, H., & Hwang, G.-H. (2024). Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technology and Society*, 215–230. https://doi.org/10.30191/ETS.202401_27(1).TP02

Cheung, B. H. H., Lau, G. K. K., Wong, G. T. C., Lee, E. Y. P., Kulkarni, D., Seow, C. S., Wong, R., & Co, M. T. (2023). ChatGPT versus human in generating medical graduate exam multiple choice questions-A multinational prospective study (Hong Kong S.A.R.,

Singapore, Ireland, and the United Kingdom). *PloS one*, 18(8), e0290691. https://doi.org/10.1371/journal.pone.0290691

Cheung, L. T., Maranna, S., Parange, N., Osborne, B., & Vu, H. (2025, December 22). Scoping review search strategy and data extraction: The Integration of Generative Artificial Intelligence into Tertiary Health Education. https://doi.org/10.17605/OSF.IO/GD3PV

Chiu, J., Castro, B., Ballard, I., Nelson, K., Zarutskie, P., Olaiya, O. K., Song, D., & Zhao, Y. (2025). Exploration of the role of ChatGPT in teaching communication skills for medical students: A pilot study. *Medical Science Educator*. https://doi.org/10.1007/s40670-025-02394-9

Çiçek, F. E., Ülker, M., Özer, M., & Kıyak, Y. S. (2024). ChatGPT versus expert feedback on clinical reasoning questions and their effect on learning: A randomized controlled trial. *Postgraduate Medical Journal*. https://doi.org/10.1093/postmj/qgae170

Collins, B. R., Black, E. W., & Rarey, K. E. (2024). Introducing anatomyGPT: a customized artificial intelligence application for anatomical sciences education. *Clinical Anatomy.* https://doi.org/10.1002/ca.24178

Cordero, J., Torres-Zambrano, J., & Cordero-Catillo, A. (2024). Integration of generative artificail intelligence in higher education: Best practices. *Education Sciences*, 15(1), 32-32. https://doi.org/10.3390/educsci15010032

Covidence systematic review software. (2024). Veritas Health Innovation. https://www.covidence.org. 2024

Crompton, H., Burke, D., Nickel, C., Bozkurt, A., Miao, F., Sharples, M., Greene, J. A., Parsons, D., Gill-Simmen, L., Edmett, A., Pegrum, M., de Waard, I., Bonk, C. J., Garcia, M. B., Curry, J. H., Lindsey, L., Yang, M., Marshall, S., Bali, M., … Yu, S. (2026). Governing generative AI in higher education: A global delphi study on policy and practice. *International Journal of Educational Technology in Higher Education*, 23(1). https://doi.org/10.1186/s41239-026-00602-z

Daniel, K., Msambwa, M. M., & Wen, Z. (2025). Can generative AI revolutionise academic skills development in higher education? A systematic literature review. *European Journal of Education*, 60(1). https://doi.org/10.1111/ejed.70036

Daniel, A. R., Nivethitha, V., Surya, B., Antoinette, D., & Visalakshi, R. (2026). Integrating ChatGPT into knowledge-retrieval tutorials in undergraduate medical education: A prospective evaluation of higher-order learning and feasibility. *Medical Education Online*, 31(1). https://doi.org/10.1080/10872981.2026.2639203

Eddine, R. J., Gide, E., & Al-Sabbagh, A. (2025). Generative AI in higher education: A cross-sector analysis of ChatGPT's impact on STEM, social sciences, and healthcare. *STEM Education*, 5(5), 757–801. https://doi.org/10.3934/steme.2025035

Elsamanoudy, A., Alhazimi, A., Hassanien, M., Almarabheh, A., Al-Aqidi, S., Alwalidi, R., & Alrefaie, Z. (2026). Implementation and evaluation of customized artificial intelligence personal tutor (chat GPT physiology companion) for medical students in the reproduction module. *BMC Medical Education*. https://doi.org/10.1186/s12909-026-09088-9

Farazouli, A., Pargman, T. C., Laksov, K. B., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment and Evaluation in Higher Education*, 49(3), 1–13. https://doi.org/10.1080/02602938.2023.2241676

Fui-Hoon Nah F. Zheng R, Cai J, Siau K, Chen L. (2023). Generative AI and ChatGPT: applications, challenges, and AI-human collaboration. *Journal of information technology case and application research,* 25(3):277-304. https://doi.org/10.1080/15228053.2023.2233814

Gao, L., Zhang, L., Zhang, Y., Wu, M., Xu, J., & Lu, D. (2026). Utilizing generative artificial intelligence to create simulated patient for history taking in gynecology. *BMC Medical Education*, 26(1). https://doi.org/10.1186/s12909-026-09219-2

Gill, S., & Galletly OAM, C. (2025). 'I, robot, can help you.' applications of generative artificial intelligence in RANZCP psychiatry training. *Australasian Psychiatry*, 33(4), 689–692. https://doi.org/10.1177/10398562251344470

Glauberman, G., Ito-Fujita, A., Katz, S., & Callahan, J. (2023). Artificial intelligence in nursing education: opportunities and challenges. *Hawai'i Journal of Health and Social Welfare,* 82(12), 302–305. https://www.ncbi.nlm.nih.gov/pubmed/38093763

Gupta, B. L., Mufti, T., Sohail, S. S., & Madsen, D. Ø. (2023). ChatGPT: A brief narrative review. *Cogent Business & Management*, 10(3). https://doi.org/10.1080/23311975.2023.2275851

Hardie, P., Darley, A., Derwin, R., Eustace-Cook, J., Kearns, S., Mc Brien, B., Siddiquee, A., Zheng, D., & Mooney, M. (2026). Applications, attitudes and ethical considerations of Generative Artificial Intelligence (Gen AI) in nursing education: a scoping review. *BMC Nursing*, 25(1), 148. https://doi.org/10.1186/s12912-025-04253-9

Hinsche, L., Hasseler, M., Tischendorf, T., & Schaal, T. (2026). Rapid review on GenAI in nursing education. *Frontiers in Health Services, 5*, 1725425. https://doi.org/10.3389/frhs.2025.1725425

Holderried, F., Stegemann-Philipps, C., Herrmann-Werner, A., Festl-Wietek, T., Holderried, M., Eickhoff, C., & Mahling, M. (2024). A language model-powered simulated patient with automated feedback for history taking: Prospective study (Preprint). *JMIR Medical Education*, 10, e59213–e59213. https://doi.org/10.2196/59213

Huang, S., Wen, C., Bai, X., Li, S., Wang, S., Wang, X., & Yang, D. (2025). Exploring the application capability of ChatGPT as an instructor in Skills education for Dental medical

students: Randomized controlled trial. *Journal of Medical Internet Research*, 27, e68538–e68538. https://doi.org/10.2196/68538

Koh, E., & Doroudi, S. (2023). Learning, teaching, and assessment with generative artificial intelligence: towards a plateau of productivity. *Learning*, 9(2), 1–8. https://doi.org/10.1080/23735082.2023.2264086

Koontz, M., & Podlog, S. (2025). Integrating generative artificial intelligence in midwifery education: Balancing innovation, ethics, and academic integrity. *Journal of Midwifery & Women's Health*. https://doi.org/10.1111/jmwh.70021

Kotzki, S., Massonnet Turner, C., Gauthier, K., Minoves, M., & Vuillerme, N. (2026). Health professional students' use of generative artificial intelligence during clinical placements: Cross-Sectional online survey study. *JMIR Medical Education*, 12, e85243–e85243. https://doi.org/10.2196/85243

Kovalainen, T., Pramila-Savukoski, S., Kuivila, H.-M., Juntunen, J., Jarva, E., Rasi, M., & Mikkonen, K. (2025). Utilising artificial intelligence in developing education of health sciences higher education: An umbrella review of reviews. *Nurse Education Today*, 147, 106600. https://doi.org/10.1016/j.nedt.2025.106600

Laye, M. J., & Wells, M. B. (2024). Rapid creation of knowledge-balanced student groups using ChatGPT4. *Medical Science Educator*, 34(3), 523–525. https://doi.org/10.1007/s40670-024-02032-w

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: a study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6(100221), 100221. https://doi.org/10.1016/j.caeai.2024.100221

Li, J., Zong, H., Wu, E., Wu, R., Peng, Z., Zhao, J., Yang, L., Xie, H., & Shen, B. (2024). Exploring the potential of artificial intelligence to enhance the writing of English academic papers by non-native English-speaking medical students - the educational application of ChatGPT. *BMC Medical Education*, 24(1). https://doi.org/10.1186/s12909-024-05738-y

Lin, Y., Luo, Z., Ye, Z., Zhong, N., Zhao, L., Zhang, L., Li, X., Chen, Z., & Chen, Y. (2025). Applications, challenges and prospects of generative artificial intelligence empowering medical education: A scoping review (preprint). *JMIR Medical Education*. https://doi.org/10.2196/71125

Lye, C. Y., & Lim, L. (2024). Generative artificial intelligence in tertiary education: assessment redesign principles and considerations. *Education Sciences*, 14(6), 569. https://doi.org/10.3390/educsci14060569

Naamati-Schneider, L. (2024). Enhancing AI competence in health management: students' experiences with ChatGPT as a learning Tool. *BMC Medical Education*, 24(1). https://doi.org/10.1186/s12909-024-05595-9

Nah, F. F.-H., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(3), 277–304. https://doi.org/10.1080/15228053.2023.2233814

NHMRC. (2018). *National Statement on Ethical Conduct in Human Research* (2007) - Updated 2018 | NHMRC. Www.nhmrc.gov.au; National Health and Medical Research Council. https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018#block-views-block-file-attachments-content-block-1

Peters M.D.J., Godfrey, C., Mclnerney, P., Munn, Z., Tricco, A., & Khalil, H. Chapter 10: scoping reviews (2020 version). In Aromataris, E., Lockwood, C., Porritt, K., Pilla, B., & Jordan, Z. editors, *JBI Evidence Synthesis*. 2020 JBI. https://doi.org/10.46658/JBIMES-24-09

Pham, T. D., Karunaratne, N., Exintaris, B., Liu, D., Lay, T., Yuriev, E., & Lim, A. (2025). The impact of generative AI on health professional education: A systematic review in the context of student learning. *Medical Educa*tion. https://doi.org/10.1111/medu.15746

Prachnakorn, N., Preecha, K., Sri-u-thai, T., Jaroenyod, T., Sawang, K., Patwong, N., & Wattanapisit, A. (2024). Incorporating artificial intelligence into a workshop on scientific and scholarly report writing for preclinical medical students. *Medical Teach*er, 46(6), 746–748. https://doi.org/10.1080/0142159x.2024.2311267

Preiksaitis, C., & Rose, C. (2023). Generative artificial intelligence in medical education: opportunities, challenges, and future directions – a scoping review (preprint). *JMIR Medical Education*, 9(1), e48785–e48785. https://doi.org/10.2196/48785

Robleto, E., Habashi, A., Kaplan, M.-A. B., Riley, R. L., Zhang, C., Bianchi, L., & Shehadeh, L. A. (2024). Medical students' perceptions of an artificial intelligence (AI) assisted diagnosing program. *Medical Teacher*, 1–7. https://doi.org/10.1080/0142159X.2024.2305369

Shoja, M. M., Van de Ridder, J. M. M., & Rajput, V. (2023). The emerging role of generative artificial intelligence in medical education, research, and practice. *Cureus*, 15(6). https://doi.org/10.7759/cureus.40883

Song, D., Zhang, P., Zhu, Y., Qi, S., Yang, Y., Gong, L., & Zhou, L. (2025). Effects of Generative artificial intelligence on Higher-Order thinking skills and artificial intelligence literacy in Nursing undergraduates: A Quasi-Experimental study. *Nurse Education in Practice*, 88, 104549–104549. https://doi.org/10.1016/j.nepr.2025.104549

Sridharan, K., & Sequeira, R. P. (2024). Artificial intelligence and medical education: application in classroom instruction and student assessment using a pharmacology & therapeutics case study. *BMC Medical Education*, 24(1). https://doi.org/10.1186/s12909-024-05365-7

Su, J., & Yang, W. (2023). Unlocking the power of chatgpt: a framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355–366. https://doi.org/10.1177/20965311231168423

Takhdat, K., El fadely, A., Mohamed, E., Ouaamr, A., & El Adib, A. R. (2026). A systematic review of generative artificial intelligence-powered healthcare simulation for clinical reasoning skills development: Applications, outcomes and challenges. *Medical Science Educator*. https://doi.org/10.1007/s40670-026-02804-6

Thomae, A. V., Witt, C. M., & Barth, J. (2024). Integration of ChatGPT Into a course for medical students: explorative study on teaching scenarios, students' perception, and applications. *JMIR Medical Education*, 10, e50545–e50545. https://doi.org/10.2196/50545

Toohey, K., Quince, Z., Walker, F., Furness, L., Bissett, M., Daley, C., Allen, K., Munro, N., Smidt, A., Wilkie, J. C., Horstmanshof, L., Baltrotsky, K., & Naumann, F. (2025). Integrating generative AI in Health education: A scoping review and implementation framework. *Medical Science Educator*. https://doi.org/10.1007/s40670-025-02578-3

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., & Lewin, S. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine*, 169(7), 467–473. https://doi.org/10.7326/M18-0850

Uribe, S. E., Maldupa, I., & Schwendicke, F. (2025). Integrating Generative AI in Dental Education: A Scoping Review of Current Practices and Recommendations. *European journal of dental education: official journal of the Association for Dental Education in Europe*, 29(2), 341–355. https://doi.org/10.1111/eje.13074

Wu, C., Chen, L., Han, M., Li, Z., Yang, N., & Yu, C. (2024). Application of ChatGPT-based blended medical teaching in clinical education of hepatobiliary surgery. *Medical Teach*er, 47(3), 445–449. https://doi.org/10.1080/0142159x.2024.2339412

Xu, X., Chen, Y., & Miao, J. (2024). Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: a systematic scoping review. *Journal of Educational Evaluation of Health Professions*, 21, 6–6. https://doi.org/10.3352/jeehp.2024.21.6

Appendices and Supplemental Materials

Appendix A

Search Strings for Google Scholar and SCOPUS In the Second Search of July 2026.

Database/ website	Search string
Google Scholar	Find articles with all the words: Integration of Generative artificial intelligence in tertiary health education, generative OR tertiary OR with the exact phrase: Generative artificial intelligence with at least one of the words: health education or medical education or chatgpt
SCOPUS	(TITLE-ABS-KEY ("Artificial Intelligence" OR "Generative AI" OR "Generative Artificial Intelligence" OR "Machine learning" OR chatgpt OR chatbot OR "Large Language Model")) AND (TITLE-ABS-KEY ("Health Educat*" OR "Medical Educat*" OR "Healthcare Student*" OR "Medical Student*" OR "Allied Health Student*" OR "Allied Health Educat*")) AND (TITLE-ABS-KEY ("Tertiary Educat*" OR undergrad* OR postgrad* OR "Higher education" OR universit*)) AND PUBYEAR > 2024 AND PUBYEAR < current