The role of generative artificial intelligence in psychiatric education– a scoping review

Qin Yuan Lee[1], Michelle Chen[1], Chi Wei Ong[1] and Cyrus Su Hui Ho[2*]

Abstract

Background The growing prevalence of mental health conditions, worsened by the COVID-19 pandemic, highlights the urgent need for enhanced psychiatric education. The distinctive nature of psychiatry– which is heavily centred on communication skills, interpersonal skills, and interviewing techniques– indicates a necessity for further research into the use of GenAI in psychiatric education.

Objective Given GenAI has shown promising outcomes in medical education, this study aims to discuss the possible roles of GenAI in psychiatric education.

Methods We conducted a scoping review to identify the role of GenAI in psychiatric education based on the educational framework of the Canadian Medical Education Directives for Specialists (CanMEDS).

Results Of the 12,594 papers identified, five studies met the inclusion criteria, revealing key roles for GenAI in case-based learning, simulation, content synthesis, and assessments. Despite these promising applications, limitations such as content accuracy, biases, and concerns regarding security and privacy were highlighted.

Conclusions Despite these promising applications, limitations such as content accuracy, biases, and concerns regarding security and privacy have been highlighted. This study contributes to understanding how GenAI can enhance psychiatric education and suggests future research directions to refine its use in training medical students and primary care physicians. GenAI has significant potential to address the growing demand for mental health professionals, provided its limitations are carefully managed.

Keywords Generative AI, Psychiatric education, Case-based learning, Simulation, Content synthesis, Assessment, CanMEDS

*Correspondence:
Cyrus Su Hui Ho
pcmhsh@nus.edu.sg
[1]School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, Singapore, Singapore
[2]Department of Psychological Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Lee *et al. BMC Medical Education*   (2025) 25:438

Page 2 of 12

Background

Generative artificial intelligence (GenAI) emulates human creativity and intelligence in the form of texts, images, videos, codes, and other modalities. According to Samala et al. (2024), it offers advantages such as cost-effectiveness, multilingual support, and efficiency [1]. It is important that educators learn to use GenAI to improve education through streamlining the process of generating educational resources and creating creative lesson plans, case-based scenarios, and assessments to deepen learners' cognitive processes.

The need for improved psychiatric education has become increasingly evident as mental health issues continue to rise globally. This rise is attributed to various factors, with the most significant being the COVID-19 pandemic, which triggered a 25% increase in the prevalence of anxiety and depression [2]. In response, some countries, such as Singapore, have extended mental health education to primary care physicians, underscoring the need to emphasise psychiatric education more [3]. However, current psychiatric education faces several challenges, including inadequate exposure to diverse patient experiences and limited resources for comprehensive training [4]. The introduction of GenAI may bridge these gaps and better prepare medical students, primary care physicians, and practitioners from other disciplines who are eager to pursue formal psychiatric education for future encounters with patients experiencing mental health-related issues.

GenAI applications in medicine can be categorised into two groups: clinical use and educational use. The clinical application of GenAI has been integrated into disease detection, diagnosis, and screening across various fields, such as radiology, cardiology, and gastrointestinal medicine [5]. GenAI has shown promising results in medical education in several areas, including self-directed learning and simulation [6].

In psychiatry, studies on the utility of GenAI primarily focus on clinical applications rather than educational purposes, such as its potential to provide diagnostic assistance, treatment considerations, and enhanced access to mental health support [7]. However, the question of whether GenAI can effectively support psychiatric education, given the unique nature of the field, has not been thoroughly addressed. The skills required of a psychiatrist place a greater emphasis on soft interpersonal skills than procedural skills, marking a significant difference from other specialities such as surgery, radiology, and endocrinology. Psychiatrists must not only be familiarised with the diagnostic criteria and prescribe appropriate medications, but they also need to master interviewing techniques and psychotherapy comprehensively while grasping phenomenology and patients' subjective experiences to formulate effective treatment plans

[8]. Many elements of psychiatric practice rely on soft skills, including conducting a Mental State Examination, suicide risk assessment, motivational interviewing, and Cognitive Behavioural Therapy. Soft skills are often more challenging to teach and evaluate than technical skills, underscoring the distinctive nature of psychiatric education [9]. This indicates that the application of GenAI in psychiatric education may differ significantly from its use in other specialities; prior studies on GenAI in medical education broadly may not be directly applicable to psychiatry.

Moreover, there is a lack of standardised guidelines regarding the use of GenAI in psychiatric education and the management of sensitive patient information and data privacy. Furthermore, GenAI may find it challenging to replicate the nuanced clinical judgement inherent in psychiatry, which heightens concerns about its accuracy. Evidence regarding the effectiveness of GenAI in enhancing psychiatric education is also limited.

By conducting a scoping review, we aim to explore our research topic by identifying GenAI's educational aspects, the benefits and risks associated with its use in psychiatric education, and the need for future research in specific areas.

Methodology

We conducted a review, limited to English publications from four databases, to identify GenAI's role in psychiatric education according to the educational framework proposed by the World Psychiatric Association-Asian Journal of Psychiatry Commission [10].

The scoping review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [11]. Our findings are presented in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) checklist [12]. A literature search in the PubMed, PsycINFO, and Embase databases was performed on 12 September 2024, followed by a search in Web of Science on 16 February 2025. A fourth database was added due to the limited number of eligible papers available for review from the first three databases. We employed the following search strategy: ("Artificial intelligence" OR "Computer reasoning" OR "Machine intelligence" OR "Machine learning" OR "Deep learning" OR "Foundation model" OR "Chat-GPT" OR "Generative AI") AND ("Mental health" OR "Psych*" OR "Psychiatric") AND ("Education" OR "Educational" OR "Training" OR "Learning" OR "Teaching"). We limited our search to English publications containing the keywords from the search strategy. The publication years for the identified papers range from 1933 to 2024.

Inclusion criteria

Original research discussing the use of GenAI or Chat-GPT in medical education was selected for review. Only original articles in English are included.

Exclusion criteria

We excluded papers discussing the clinical use of GenAI, public/patient mental health education, technology in general, virtual reality, and augmented reality, as well as papers addressing a specific field of medical education that is not related to psychiatry (e.g., oncology, surgical), nursing, psychology, and the perception of GenAI. We also excluded conference papers, preprints, editorials, and other non-original research.

Studies selection process

Search results from the databases were uploaded to End-Note. Duplicates were removed, followed by title and abstract screening using the inclusion and exclusion criteria. LQY and MC conducted the initial screening independently. Studies deemed eligible were downloaded, and full-text screening was carried out by LQY and MC. Any disagreements were resolved by consulting a third senior reviewer (CWO and CSH).

Data extraction and analysis

Details of the reviewed paper, such as authors, year of publication, type of GenAI, methodology, outcome measure, and key findings, were charted in a table by LQY and MC (refer to Table 1). Through thematic analysis, the role of GenAI in psychiatric education was grouped into four themes, and evidence synthesis was done to achieve the aim of this study. The senior authors checked the tabulation of data, themes, and syntheses.

Results

We identified 12,594 papers, of which 118 were duplicates. After the abstract screening, 12,439 papers were excluded. 37 papers were reviewed in full text, and 32 papers were excluded because they did not meet the inclusive criteria (refer to Fig. 1). The remaining 5 papers discussing using GenAI in medical education were selected for review.

The types of GenAI used include ChatGPT (3.5 and 4), Claude 3, and Llama 3. All papers addressed the use of ChatGPT. Most examined the differences between content generated by GenAI and that produced by traditional handwritten or expert-written sources. The five papers explored four roles that GenAI can fulfil in medical education: case-based learning, simulation, content synthesis, and assessments [7, 13–16].

Case-based learning

Two papers discussed leveraging GenAI to create case vignettes for case-based learning [7, 13]. In a study by Coşkun et al. (2024), a randomised controlled trial was conducted to compare the quality of ChatGPT-synthesised vignettes with those written by humans. There was no significant difference in quality between the two types of vignettes. The scores suggested that vignettes generated by ChatGPT may promote higher utilisation of clinical reasoning skills among students compared to those created by humans.

Furthermore, the study by Smith et al. (2023) highlighted the efficiency and variety of ChatGPT-generated case vignettes. Educators can modify these vignettes to teach various learning outcomes, such as the diagnostic process, treatment, determining if a psychopharmacological therapy is necessary, or the ethics surrounding the case. For instance, adjusting the prompt to exclude suggestions for treatment plans allows students to discuss and describe the factors to consider when prescribing a treatment [7]. The parameters of the case can also be varied, including its difficulty and complexities, offering flexibility in the creation of course materials. Psychiatric disorders often present a complicated array of symptoms. A diverse range of case vignettes could better prepare medical students to diagnose and treat patients.

Other advantages include addressing ethical concerns associated with utilising real case vignettes and the capacity to produce case vignettes in various languages [7]. Real case vignettes must undergo rigorous scrutiny and documentation to guarantee informed consent and maintain patient confidentiality. This is particularly challenging in psychiatry, as patients may lack the mental capacity to consent, and ensuring anonymity can be problematic [17]. GenAI-generated vignettes do not face the same issues.

Simulation

ChatGPT is recognised for its ability to conduct simulations by adopting character roles and providing real-time responses based on input [1]. One study by Smith et al. (2023) briefly noted that ChatGPT could simulate a patient, facilitating interactions with students to practise their clinical skills or their ability to identify risk factors [7]. Previously, a review indicated that simulation in psychiatry effectively enhances students' competencies in performing psychiatric risk assessments on patients [18]. However, a shortage of studies exists that address the methods and effectiveness of GenAI in patient simulation within psychiatric education, which would facilitate its implementation.

Lee *et al. BMC Medical Education*   (2025) 25:438

Page 4 of 12

Table 1 Selected characteristics of included studies

Author	Year of Publication	Aim	Type of GenAI	Themes	Methodology	Outcome measure	Key Findings
Smith et al. [7]	2023	Demonstrate the possible ways that ChatGPT could be utilised to aid education, specifically for social psychiatry.	ChatGPT 3.5	Case-based learning; Simulation	ChatGPT was prompted to highlight its functions in aiding social psychiatric education and demonstrate one of them. A case vignette was generated based on a prompt.	Comparison made between ChatGPT-generated vignettes and educational materials from the Department of Forensic Psychiatry at the University of Bern.	There are many ways GenAI can contribute to psychiatric education. GenAI can circumvent impediments of traditional education methods.
Coşkun et al. [13]	2024	Assess the performance of ChatGPT generated clinical vignettes and MCQs in practical situations.	ChatGPT 3.5	Case-based learning	Medical students evaluated either Chat-GPT generated or human-written vignettes.	Likert-scale scoring of 10 questions by medical students regarding the ChatGPT or handwritten vignettes.	Quality of ChatGPT-generated vignettes are comparable to human-written.
				Assessment	Medical students answered 15 ChatGPT-generated MCQs.	Psychometric characteristics of MCQs.	Some MCQs generated by ChatGPT have acceptable psychometric characteristics.
Hudon et al. [14]	2024	Compare ChatGPT-generated SCT and those written by clinical experts in undergraduate psychiatric education.	ChatGPT 3.5	Assessment	3 SCTS generated from ChatGPT and 3 created by experts were evaluated by clinician educators.	Qualitative analysis of the advantages and disadvantages of Chat-GPT vignettes. Statistical analysis of whether case vignettes were correctly identified as ChatGPT generated or handwritten.	There is no major difference between ChatGPT-generated SCTs and those created by experts. The caricatural quality and significance of SCTS generated by Chat-GPT could be further improved.
Kiyak et al. [15]	2024	Develop prompts to generate Script Concordance Test items, a common assessment tool in medical education, using large language models chatbots.	ChatGPT-4 Claude 3 Llama 3 ChatGPT-4o	Assessment	A prompt was created using a prompt template and used in 4 chatbots.	Ease of generating SCTs.	GenAI, with appropriate prompts, can streamline the process of generating SCT items.
Yanagita et al. [16]	2024	The feasibility of using generative AI to create illness scripts.	ChatGPT-4	Content synthesis and summary	Illness scripts were generated for 184 diseases. The scripts were evaluated and graded by 3 physicians.	Illness scripts were ranked by "A," "B," "C" where A shows that the script is the most informative and suitable script to be used by medical students.	More than half the illness scripts were rated "A," which is the highest tier and required no further edits to them.
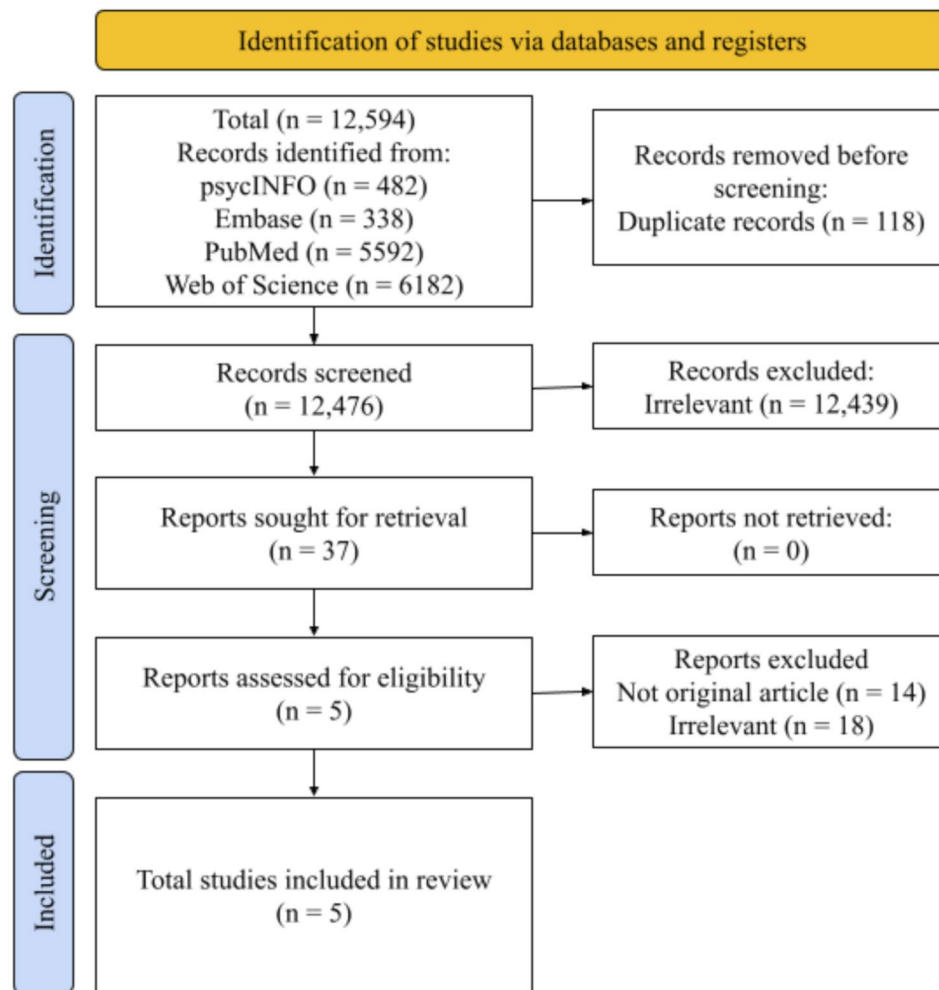
Fig. 1 Flow chart showing identification, screening and inclusion of papers reviewed according to PRISMA [11]

Content synthesis and summary

ChatGPT can streamline the content synthesis process, enhancing efficiency while upholding academic standards [1]. It has been shown to provide accurate medical information and simplified summaries of complex research [7]. Specifically, a paper discusses using GenAI to create illness scripts for educational purposes [16]. An illness script is a specific format representing patient-oriented clinical knowledge containing valuable information. It is generally dynamic, depending on the physician's requirements, but it can also be standardised for medical education. Illness scripts can teach medical students clinical reasoning skills, thereby improving diagnostic accuracy. In the study by Yanagita et al. (2024), 84% of the 184 illness scripts demonstrated relatively high accuracy.

Assessment

Three papers explored the application of GenAI to develop assessment tools for medical students [13–15].

Coşkun et al. (2024) discussed the quality of ChatGPT-generated multiple-choice questions (MCQs). Out of 15 questions generated, six items met the criteria and were concluded to be effective. The quality of MCQs can be further refined by using more complex prompts, including factors such as learner type, competency level, and difficulty level [13].

In addition to generating MCQs, two papers discussed the creation of the Script Concordance Test (SCT) using Large Language Models (LLMs) [14, 15]. The SCT is designed to refine clinical reasoning and decision-making in uncertain clinical situations. Developing an SCT is both challenging and complex; therefore, employing GenAI-like LLMs can help expedite the development process [15]. Hudon et al. (2024) examined the application of ChatGPT-generated SCTs in psychiatry for undergraduate medical education. It has been demonstrated that there is no significant difference between ChatGPT-generated SCTs and those created by experts in terms of the scenario, clinical questions, and expert

opinions. With the appropriate target group, a relevant focus of the question, the clinical problem, and guidelines to follow, an SCT for psychiatric education can be easily developed. This can be conveniently achieved using the "Script Concordance Test Generator," a custom GPT designed for SCT generation [15].

Overall limitations of GenAI

Overall, there are concerns regarding inaccuracies, bias, and a lack of control over generated content [7, 13]. Moreover, using GenAI for simulation poses the risk of sharing sensitive or personal data, thereby raising security and privacy issues [7]. ChatGPT may also display grammatical errors in certain languages, exhibit biases against minorities, experience hallucination effects, show a lack of replicability, possess limited awareness of recent events, and may eventually adopt a paywall, leading to inequality [7].

Moreover, GenAI-generated illness scripts for psychiatric disorders received the highest number of "C" ratings, comprising 45.5% of psychiatric scripts [16]. These scripts present generic information, such as "diagnosis based primarily on clinical interview and symptom criteria", instead of outlining the specific steps involved [16]. This issue may arise from the limited character count; however, this constraint could be alleviated by increasing it. With a greater character count, more details could be covered, particularly considering the wide variety of psychiatric symptoms.

Another limitation of GenAI is that the SCTs it generated were too simple [14]. Well-designed and more complex prompts can improve the quality of SCTs. Subject matter experts can make minor adjustments. Appropriate guidelines are still needed to leverage GenAI, as the content generated may not meet the standard required for education use. For example, "none of the above" is discouraged by test development guidelines for MCQ, but it was included as an option for one of the generated questions [13].

Discussion

Despite the limited number of papers, GenAI has demonstrated its potential role in psychiatric education. While the role of GenAI is extensively discussed in other specialities and for clinical applications, there is minimal analysis regarding its use in psychiatric education. The intricate nature of psychiatry may be one factor contributing to the lack of exploration into the role of GenAI in this field [19]. In this section, we will utilise an established psychiatric education framework to analyse the applicability of GenAI's role in psychiatric education.

This review demonstrates favourable evidence that GenAI, such as ChatGPT, supports psychiatric education through case-based learning, simulation, content

analysis, and assessment. To explore the potential of GenAI in psychiatric education, we compare our findings with the new training framework established by the World Psychiatric Association-Asian Journal of Psychiatry Commission. This framework is based on The Canadian Medical Education Directives for Specialists (CanMEDS) developed by The Royal College of Physicians and Surgeons of Canada in response to the evolving landscape of psychiatry [10]. CanMEDS is applicable to various disciplines, including psychiatric education [20]. It comprises seven competencies: communicator, collaborator, leader, health advocate, scholar, professional, and medical expert (see Fig. 2). The World Psychiatric Association-Asian Journal of Psychiatry Commission paper discussed specific requirements and recommendations necessary to achieve these seven competencies given Psychiatry's increasingly complex and uncertain nature (see Table 2). Therefore, by analysing how GenAI can contribute to this framework, we can provide improved solutions for delivering quality psychiatric education that meets modern needs. In the following sections, we explore how the use of GenAI in case-based learning, simulation, content analysis, and assessment contribute to shaping the seven competencies of CanMEDS.

Case-based learning

Generative AI opens up opportunities for creating various case vignettes. The effectiveness of case-based learning in psychiatric training has been highlighted in previous research [21]. An efficient method for generating case vignettes can contribute to the development of roles such as medical expert, communicator, collaborator, leader, scholar, and professional.

Moreover, case vignettes can be adapted to teach the use of diagnostic tools and criteria while practising within legal frameworks and safety protocols. Most importantly, the ability to vary case vignettes can train medical students to handle situations involving uncertainty or dilemmas. This is crucial in today's world, where mental health issues are complicated by contemporary factors such as social media influences and non-evidence-based self-diagnostic tools found online [22]. In this context, GenAI can potentially support the role of a medical expert. Understanding how to respond to cases involving dilemmas also enhances the roles of communicator, leader, and professional.

Furthermore, GenAI can be prompted to create cases that necessitate interdisciplinary collaboration (e.g., a combination of mental and physical illnesses or a scenario where the involvement of a financial adviser or social worker is essential). This fosters medical students' development of collaborative skills. During discussions of the cases, medical students can explore how to integrate
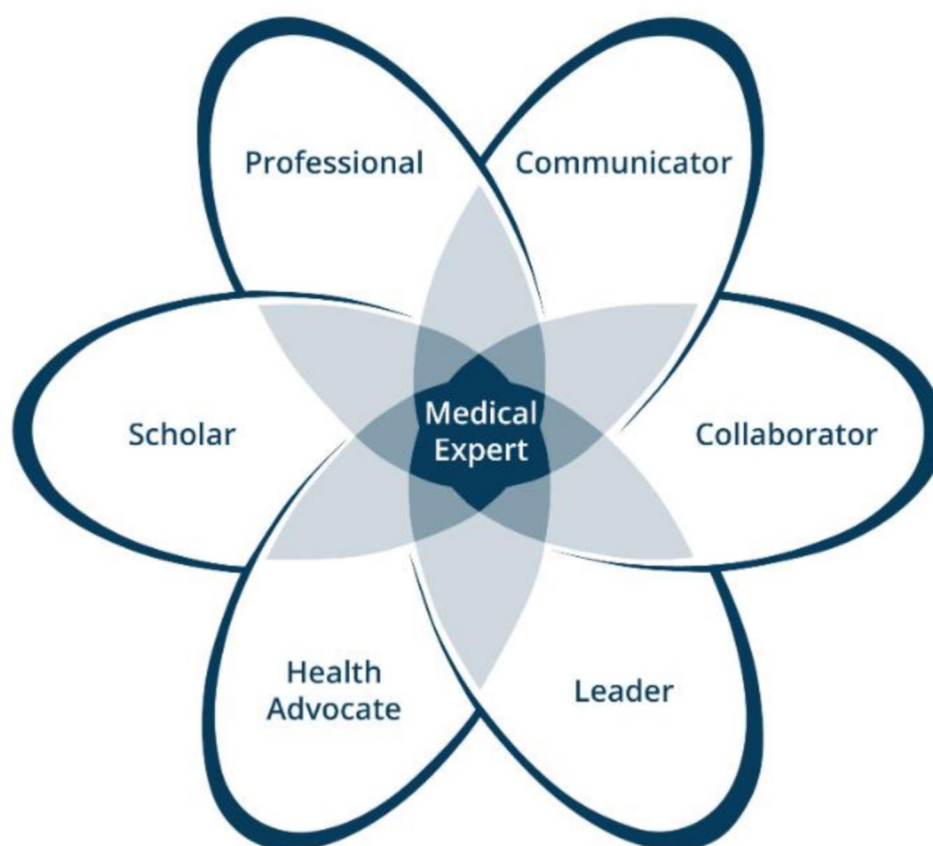
Fig. 2 CanMEDS Diagram designed by the Royal College of Physicians and Surgeons of Canada [10]

evidence-based knowledge alongside patients' values and preferences. This competency is expected of a scholar.

Simulation

In addition to case-based learning, applying what students learn in practice is essential in psychiatry. Compared to other specialities, communication in psychiatry must be extremely precise and sensitive to patients. By having GenAI simulate patient dialogues, students can practise the communication framework they have learnt and learn to be flexible with these communication techniques, such as motivational interviewing for addiction. This aligns with the roles of medical expert and communicator. Students can also learn to maintain professional boundaries by carefully selecting their words during conversations with the simulator, aiding their development into professionals.

Through simulation, students can better convey information regarding treatments, apply diagnostic assessments, and gather comprehensive patient histories. However, Dave (2012) highlighted several concerning limitations associated with implementing simulated patients in psychiatric education. Given that mental health illnesses are often complicated to understand, it is challenging to train simulated patients to accurately

portray the complexities of psychiatric conditions [23]. Additionally, actors may introduce and act upon their prejudices towards mental illness. Cost is also a significant concern. Interestingly, some studies discuss the use of GenAI-based 2D or 3D avatars to enhance patient encounters in other specialities [24]. GenAI-based simulators could assist in overcoming these challenges, provided there are no inherent biases or paywalls. Further research into its application in psychiatric education is warranted.

Content synthesis and summary

To encourage medical students to embody the role of medical experts and scholars, GenAI synthesises illness scripts, enabling students to grasp essential information regarding various diseases. However, given the complexities of psychiatric illnesses, further studies are necessary to enhance the quality and examine the effectiveness of GenAI-created illness scripts in psychiatric education.

Furthermore, GenAI can also promote lifelong learning, provided graduated healthcare workers are granted free access to it, or no paywalls are implemented in the future. This allows them to obtain the information as and when it becomes available. However, as highlighted in the

Table 2 Summary of CanMEDS and recommendations by the world psychiatric Association-Asian journal of psychiatry commission [10]

	Requirements	Recommendations
Medical expert	• Able to recognise both mental and somatic conditions • Able to understand diagnostic criteria, including around co-morbidities or complex conditions for more comprehensive care • Able to use standardised assessment tools • Able to conduct thorough patient assessment • Able to use all available information to accurately diagnose • Able to provide appropriate evidence-based treatment • Able to explain information about the treatments • Able to practise within the legal framework and safety protocols • Learn to deal with clinical uncertainty • Learn the biopsychosocial model	• Psychiatric training should be well-integrated into undergraduate education • Inclusion of non-traditional teaching materials • Simulation • Reflective practice • Clinical vignettes • (role) modelling • Role playing • Group teaching with patients • Using videotapes • Bedside teaching • Research evaluation
Communicator	• Be aware of roles of different team members and intra-team communication. • interpersonal and relationship management skills, such as active listening and nonverbal communication • Able to utilise flexible techniques • Allow shared-decision making and informed decision • Able to handle dilemmas • Promote patient's well-being	• Observing role model • theoretical teaching • Video feedback • Virtual reality and avatars • Directly including service users in interpersonal training programmes • Simulation • Cinema or cultural productions • Interacting with patients' families at discharge meetings. • Cultural competency training • Self-assessment • Group-based discussion • Shared reflection
Collaborator	• Ability to work with different stakeholders • Collaboration beyond clinical scenario • Able to support handovers in patient care • Be aware of situations where knowledge sharing and collaboration is not appropriate	• Group-based assessments • Hands-on interprofessional approach with patient involvement • Interdisciplinary meetings • Comprehensive treatment plans under appropriate supervision • Joint learning
Leader	• Ability to engage with other stakeholders and policy maker • Ability to engage in societal outreach • Awareness of recent trends • Ability to offer integrated care • More emphasis on performance variability and risk • Sensitive approach in face of dilemmas • Lifelong learning and professional development	• Formal learning objectives • Cross-disciplinary approaches • Transformative learning
Health Advocate	• Advocate using expertise and influence • Reduce stigma for minorities • Promote mental health literacy • Removing systemic barriers • Support expansion of mental health workforce • Promote integration of mental health care into public health and social policy	• Formalised training • Advocacy praxis • Transformative learning • Teach psychiatrist to work with community leaders • Mix of theoretical and experiential teaching • Lectures on social determinants • Hands-on experiences • Didactic curricula • Mentorship • Self-assessment
Scholar	• Be equipped with latest knowledge • Able to consider patient values and preferences into account • Lifelong learning • Serve as a mentor	• Team-based learning • Group discussions • Evidence-based medicine courses • Didactic and problem-based techniques • Protected time for teaching and learning in clinical settings for trainees • Journal clubs • Grand rounds • Research groups on WhatsApp • Engaged in teaching, conference presentations, leading workshops

Table 2 (continued)

	Requirements	Recommendations
Professional	• Perform ethical, moral and professional responsibilities • Putting others needs above their own • Maintain public's trust in the profession • Uphold confidentiality • Navigate through uncertain scenarios • Emotional skills • Establish clear boundaries • Self aware of own biases • Accountability • Openness to feedback • Responsible for own well-being	• Innovative didactic pedagogies • Resources and workshops on stress and burnout • Supervision • Peer support groups • Balint groups • Mentorship • Experiential learning opportunities • Team-based roleplay

results section, GenAI has its inaccuracies. Users need to be careful and stay cautious when using GenAI.

Assessment

In the results section, we discussed using GenAI to generate MCQs and SCTs. These can either serve as summative assessments or as self-quizzes for students to prepare for an assessment.

Most medical education exams are conducted as MCQs, or at least in Singapore. Thus, students can practise applying knowledge by generating and completing MCQs for self-preparation. However, this is provided that GenAIs that can generate quality MCQs do not have paywalls; otherwise, this could facilitate greater inequality between different income groups.

The use of SCTs in psychiatry has been studied, and the feasibility of evaluating clinical reasoning has been shown [25]. SCTs can be adapted to assess whether students fulfil the roles of CanMEDS. They have the potential to assess psychiatry clinical competencies, such as understanding diagnostic frameworks and clinical assessment tools, dealing with uncertainty, and practising evidence-based medicine. These competencies are required of a medical expert, leader, and professional.

In both MCQs and SCTs, GenAI can easily generate a diverse range of questions. These questions can incorporate the socio-economic or racial backgrounds of the patient, allowing for the assessment of the student's objectivity and training them to remain non-judgemental. This could enhance the student's role as a health advocate, helping to reduce stigma for patients, particularly those from minority groups.

Incorporation of GenAI in psychiatric education

The different applications may work and integrate together—for example, GenAI–created case vignettes can be used as a prompt to generate video simulations, and GenAI can assess students' answers to GenAI-created questions. In addition to the four applications discussed in this study, other applications can be explored, such as using GenAI to translate content into various languages, demolishing language barriers, and promoting access to psychiatric education resources in more countries, contributing to global mental healthcare [1].

However, implementing GenAI in psychiatric education may present some challenges. Educators might hesitate to adopt GenAI since the current approach remains traditional and predominantly face-to-face. They may worry about the potential loss of warmth, empathy, and personal interactions from using GenAI. Many educators and clinicians are not yet trained to use GenAI tools for psychiatric education. There may also be scepticism about whether GenAI can enhance traditional case-based discussions, psychotherapy training, or diagnostic reasoning exercises.

Addressing the risks of GenAI

GenAI can produce hallucinations– the generation of factually inaccurate information [26]. AI hallucinations occur when AI creates seemingly realistic but entirely fabricated content that may be illogical or incorrect [27]. Several reasons for the occurrence of AI hallucinations include insufficient diversity in training data or biases rooted in certain background traits. GenAI-generated content, such as illness scripts, may lack accuracy [1]. Students who study these illness scripts without expert revisions risk grasping incorrect medical concepts, which could lead to poor medical decisions in the future. Similarly, Coşkun et al. (2024) highlighted that inaccurate information was identified in GenAI-generated clinical vignettes and MCQs, posing the risk of disseminating incorrect information to students [13].

As GenAI heavily relies on training data to generate outputs, assessment questions and vignettes produced by ChatGPT may follow a predictable pattern. This might result in a limited variety of exam questions, failing to encapsulate the sophistication of psychiatric education [28].

When it comes to GenAI, privacy is a significant concern. In this study, however, GenAI can assist with the issue of patient privacy by removing the use of real case vignettes for case-based learning and SCT generation.

Nevertheless, there is a risk of question banks being leaked to medical students. The outputs (e.g., generated examination questions) of GenAI may be stored in the AI system, which raises the possibility of the questions being leaked to students using the same AI system [29].

While GenAI presents several risks, including ethical concerns and inaccuracies, these issues can be effectively managed through specific recommendations. Further research should focus on establishing clearer guidelines for GenAI usage in psychiatric education founded on ethical principles. Additionally, the potential for bias in GenAI could be alleviated by training it with more comprehensive datasets. The data utilised must adhere to data protection laws. Furthermore, experts should conduct a manual review to evaluate the accuracy and relevance of GenAI-generated content. Technologies such as Federated Learning and Blockchain can be explored as potential solutions to the issue of question leaks in psychiatric education assessments.

Limitations of study

Here, we analyse the quality of the studies reviewed and identify the strengths and limitations of each study reviewed:

The study by Smith et al. (2023) examined the various applications of ChatGPT in depth; however, it lacked a definitive methodology for assessing its effectiveness in social psychiatry.

The study by Coşkun et al. (2024) is a randomised controlled experiment that employs strong methodology and psychometric evidence to justify ChatGPT's potential in generating clinical vignettes and MCQs for assessment. However, this study does not directly address psychiatric education.

Kıyak et al. (2024) examined various types of GenAI beyond ChatGPT and proposed specific prompts for generating SCT items, thereby justifying the potential to streamline the creation of complex educational materials. However, this study did not assess the effectiveness of GenAI-generated SCTs in improving psychiatric educational outcomes.

Hudon et al. (2024) have a methodology designed to avoid biased results. A considerable number of clinician-educators and resident doctors evaluated the effectiveness of ChatGPT in psychiatric education. Future studies may consider adopting their framework to assess the effectiveness of GenAI in psychiatric education in other ways.

Yanagita et al. (2024) analysed a considerable number of ChatGPT-generated illness scripts. However, only three physicians reviewed the quality of these illness scripts, amplifying the issue of subjectivity.

Considering the limitations of existing studies, future research could employ a more quantitative measure to assess GenAI's effectiveness on student outcomes, explore different types of language models, and involve a larger study size.

The risk of publication bias in selecting articles was minimised during the screening process through independent assessments and third-party opinions. However, the limited literature search yielded only two studies that directly addressed psychiatric education, while the remaining three studies focused more generally on medical education, leading to generalisations from medical education to psychiatry. The small number of relevant studies restricts the generalisability of our findings and discussion. The reviewed papers did not permit quantitative analysis and were not comparable. The absence of a quantitative, comparative analysis for drawing conclusions is a limitation of our study. Nonetheless, this underscores the need for further research in this area.

Comparison with prior work

To our knowledge, no paper has discussed the use of GenAI in psychiatric education. Prior studies mainly focused on the use of GenAI in clinical psychiatry or medical education in general but did not discuss its suitability in psychiatric education.

There are several reasons why psychiatric education has received less attention regarding the incorporation of GenAI. Firstly, the skills a psychiatrist must acquire are highly humanistic, emphasising the doctor-patient relationship [8]. Employing GenAI, essentially a non-human entity, to teach psychiatry is an unusual approach at first glance, thus making it a topic that is rarely discussed. In contrast, for other specialities such as radiology, GenAI can directly assist with technical skills important in the field—such as generating images of pathological findings (for example, x-ray imaging and skin lesions) as training materials [30]. This is not applicable in psychiatry, as the diagnosis and management of mental health disorders can be subjective and cannot be easily determined by observing images.

Secondly, since soft skills are the core competencies required of a psychiatrist, it is essential to evaluate students' performance based on these skills. GenAI may not accurately assess this, as it fundamentally lacks a deep understanding of empathy and emotional states [1]. However, in other fields, GenAI can be effectively utilised to assess performance and provide appropriate feedback. For instance, the OpenAI GPT-4 Turbo API could review revisions of radiology reports made by trainees and generate relevant educational feedback [31].

Conclusion

Our scoping review showed that Generative AI has potential in psychiatric education. GenAI can complement traditional pedagogies, gearing psychiatric

education toward achieving the goal of CanMEDS. This suggests that GenAI can cater to the unique nature of psychiatry.

Nevertheless, this area remains largely unexplored. Limitations such as content accuracy, privacy, and ethical concerns must be addressed. Further research and measures should be established before implementing GenAI. Future studies should address these limitations, propose mitigating strategies, and evaluate GenAI's effectiveness on educational outcomes and how such outcomes contribute to students' performance in clinical practice. There is a need to enhance the engagement of individuals researching the use of GenAI in psychiatric education, discover more effective methods to cater to the nature of psychiatry and encourage educators to be more receptive to participating in the research and implementation of GenAI in psychiatric education. Comprehensive studies on the cost-benefit analysis of implementation should be conducted, assessing benefits (e.g., student outcomes and educator efficiency) against costs (e.g., expenses of integrating GenAI and addressing potential ethical concerns). With further studies, potential breakthroughs in psychiatric education may be realised.

Abbreviations

GenAI	Generative Artificial Intelligence
CanMEDS	The Canadian Medical Education Directives for Specialists
SCT	Script Concordance Test
MCQ	Multiple Choice Questions

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References
1. Samala AD, Rawas S, Wang T, Reed JM, Kim J, Howard N-J, Ertz M. Unveiling the landscape of generative artificial intelligence in education: a comprehensive taxonomy of applications, challenges, and future prospects. Educ Inf Technol. 2024;1–40. https://doi.org/10.1007/s10639-024-12936-0
2. WHO. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. World Health Organ. 2022. Available from: https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide [accessed Sep 30, 2024].
3. Peh A, Tan G, Soon W, Cheah B, Ng J. Psychiatry in primary care and training: a Singapore perspective. Singap Med J. 2021;62(5):210–2. https://doi.org/10.11622/smedj.2021056
4. Sampogna G, Elkholy H, Baessler F, Coskun B, Pinto da Costa M, Ramalho R, Riese F, Fiorillo A. Undergraduate psychiatric education: current situation and way forward. BJPsych Int 19(2):34–6. PMID:35532467.
5. Yim D, Khuntia J, Parameswaran V, Meyers A. Preliminary Evidence of the Use of Generative AI in Health Care Clinical Services: Systematic Narrative Review. JMIR Med Inf. 2024;12:e52073. PMID:38506918.
6. Preiksaitis C, Rose C. Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review. JMIR Med Educ. 2023;9(1):e48785. https://doi.org/10.2196/48785
7. Smith A, Hachen S, Schleifer R, Bhugra D, Buadze A, Liebrenz M. Old dog, new tricks? Exploring the potential functionalities of ChatGPT in supporting educational methods in social psychiatry. Int J Soc Psychiatry. 2023;69(8):1882–9. https://doi.org/10.1177/00207640231178451
8. Fountoulakis KN. Psychiatry among the Other Medical Specialties., Springer P. Cham; 2022:469–476. https://doi.org/10.1007/978-3-030-86541-2_19ISBN:978-3-030-86541-2.
9. Iorio S, Cilione M, Martini M, Tofani M, Gazzaniga V. Soft Skills Are Hard Skills—A Historical Perspective. Med (Mex). 2022;58(8):1044. PMID:36013513.
10. Bhugra D, Smith A, Ventriglio A, Hermans MH, Ng R, Javed A, Chumakov E, Kar A, Ruiz R, Oquendo M. World Psychiatric Association-Asian Journal of Psychiatry Commission on psychiatric education in the 21st century. Asian J Psychiatry Elsevier. 2023;88:103739.
11. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ Br Med J Publishing Group. 2021;372:n71. PMID:33782057.
12. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MDJ, Horsley T, Weeks L, Hempel S, Akl EA, Chang C, McGowan J, Stewart L, Hartling L, Aldcroft A, Wilson MG, Garritty C, Lewin S, Godfrey CM, Macdonald MT, Langlois EV, Soares-Weiser K, Moriarty J, Clifford T, Tunçalp Ö, Straus SE. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med Am Coll Physicians. 2018;169(7):467–73. https://doi.org/10.7326/M18-0850
13. Coşkun Ö, Kıyak YS, Budakoğlu İİ. ChatGPT to generate clinical vignettes for teaching and multiple-choice questions for assessment: A randomized controlled experiment. Med Teach 2024;((Coşkun Ö.; Kıyak Y.S.; Budakoğlu İ.İ.) Department of Medical Education and Informatics, Gazi University, Ankara, Turkey):1–7. https://doi.org/10.1080/0142159X.2024.2327477
14. Hudon A, Kiepura B, Pelletier M, Phan V. Using ChatGPT in Psychiatry to Design Script Concordance Tests in Undergraduate Medical Education: Mixed Methods Study. JMIR Med Educ. 2024;10:e54067. PMID:38596832.
15. Kıyak YS, Emekli E. A Prompt for Generating Script Concordance Test Using ChatGPT, Claude, and Llama Large Language Model Chatbots. Rev Esp Educ Médica. 2024;5(3). Available from: https://revistas.um.es/edumed/article/view/612381 [accessed Sep 14, 2024].
16. Yanagita Y, Yokokawa D, Fukuzawa F, Uchida S, Uehara T, Ikusaka M. Expert assessment of ChatGPT's ability to generate illness scripts: an evaluative study. BMC Med Educ. 2024;24(1):536. https://doi.org/10.1186/s12909-024-05534-8
17. Draper H, Rogers W. Re-evaluating confidentiality: using patient information in teaching and publications. Adv Psychiatr Treat. 2005;11(2):115–21. https://doi.org/10.1192/apt.11.2.115
18. Ghahari D, Chaharlangi D, Bonato S, Sliekers S, Sockalingam S, Ali A, Benassi P. Educational approaches using simulation for psychiatric risk assessment: A scoping review. Acad Psychiatry. 2024;48(1):61–70.
19. Fried EI, Robinaugh DJ. Systems all the way down: embracing complexity in mental health research. BMC Med. 2020;18:205. PMID:32660482.
20. Tuhan I. Mastering CanMEDS Roles in Psychiatric Residency: A Resident's Perspective. Can J Psychiatry SAGE Publications Inc. 2003;48(4):222–4. https://doi.org/10.1177/070674370304800404

21. McParland M, Noble LM, Livingston G. The effectiveness of problem-based learning compared to traditional teaching in undergraduate psychiatry. Med Educ. 2004;38(8):859–67. https://doi.org/10.1111/j.1365-2929.2004.01818.x

22. Corzine A, Roy A. Inside the black mirror: current perspectives on the role of social media in mental illness self-diagnosis. Discov Psychol. 2024;4(1):40. https://doi.org/10.1007/s44202-024-00152-3

23. Dave S. Simulation in psychiatric teaching. Adv Psychiatr Treat. 2012;18(4):292–8. https://doi.org/10.1192/apt.bp.110.008482

24. Sardesai N, Russo P, Martin J, Sardesai A. Utilizing generative conversational artificial intelligence to create simulated patient encounters: a pilot study for anaesthesia training. Postgrad Med J. 2024;100(1182):237–41. https://doi.org/10.1093/postmj/qgad137

25. Kazour F, Richa S, Zoghbi M, El-Hage W, Haddad FG. Using the Script Concordance Test to Evaluate Clinical Reasoning Skills in Psychiatry. Acad Psychiatry. 2017;41(1):86–90. https://doi.org/10.1007/s40596-016-0539-6

26. Ning Y, Teixayavong S, Shang Y, Savulescu J, Nagaraj V, Miao D, Mertens M, Ting DSW, Ong JCL, Liu M, Cao J, Dunn M, Vaughan R, Ong MEH, Sung JJ-Y, Topol EJ, Liu N. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. Lancet Digit Health Elsevier. 2024;6(11):e848–56. PMID:39294061.

27. Athaluri SA, Manthena SV, Kesapragada VSRKM, Yarlagadda V, Dave T, Duddumpudi RTS. Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing Through ChatGPT References. Cureus. 2023;15(4):e37432. PMID:37182055.

28. V EKSPRGRKLABMGTOSR. Advantages and pitfalls in utilizing artificial intelligence for crafting medical examinations: a medical education pilot study with GPT-4. BMC Med Educ. 2023;23:772. PMID:37848913.

29. Diro A, Kaisar S, Saini A, Fatima S, Hiep PC, Erba F. Workplace security and privacy implications in the GenAI age: A survey. J Inf Secur Appl. 2025;89:103960. https://doi.org/10.1016/j.jisa.2024.103960

30. Janumpally R, Nanua S, Ngo A, Youens K. Generative artificial intelligence in graduate medical education. Front Med Front. 2025;11. https://doi.org/10.3389/fmed.2024.1525604

31. Lyo S, Mohan S, Hassankhani A, Noor A, Dako F, Cook T. From Revisions to Insights: Converting Radiology Report Revisions into Actionable Educational Feedback Using Generative AI Models. J Imaging Inf Med. 2024. https://doi.org/10.1007/s10278-024-01233-4

Publisher's note