Medical Science Educator
https://doi.org/10.1007/s40670-025-02473-x

ORIGINAL RESEARCH

Generative AI as the First Patient: Practice, Feedback, and Confidence

Thomas Thesen[1,2] · **Wade N. O'Brien**[3] · **Simon Stone**[4] · **Roshini Pinto-Powell**[1,5]

Accepted: 16 July 2025
© The Author(s) under exclusive licence to International Association of Medical Science Educators 2025

Abstract
Generative artificial intelligence (GenAI) offers new potential for medical education by providing personalized learning opportunities at scale that address current faculty shortages, such as limited access to clinical preceptors, yet its integration into formal medical training and acceptability by students remains largely unexplored. We deployed a large-language model-based AI Patient Actor app in a first-year medical curriculum, enabling students to conduct simulated clinical interviews while receiving immediate feedback. Students rated their experiences as overwhelmingly positive, with comfort using GenAI in general increasing significantly after repeated engagement. Sentiment analysis of student comments showed appreciation for the AI approaches' flexibility, low-pressure environment, and real-time feedback that students felt helped with clinical reasoning and communication skills. Students identified limitations regarding the authenticity of AI interactions compared to human encounters. This study shows that AI patient actors are perceived as acceptable educational adjuncts in early clinical training, particularly in formative environments where they provide students with immediate feedback and greater autonomy over their learning process, though they cannot fully replace interactions with human actors and patients.

Keywords Generative AI · AI Patient Actor · Medical curriculum

Introduction

Medical education is undergoing notable change driven by recent advances in generative AI (GenAI) technologies. While medical students demonstrate enthusiasm about AI and believe it will play a central role in their future careers, there remains limited integration of AI into current educational curricula[1–3]. Consequently, the application of GenAI in medical training represents an emerging field that is still largely unexplored, with minimal research on learner perspectives regarding its effectiveness, limitations, and optimal implementation strategies. GenAI's substantial potential for medical education lies in its ability to provide continuous, unrestricted access to learning opportunities beyond traditional classroom constraints. This technology enables individualized instruction to create experiences tailored to each student's unique needs[4]. Through conversational interfaces, large-language models (LLMs) can facilitate dynamic, interactive exchanges and feedback that develop communication and cognitive skills through interactive engagement, outcomes previously achievable only with significant faculty resources. This approach aligns with precision education frameworks, which advocate for tailoring instruction to individual learner characteristics through adaptive and data-driven approaches with a goal of improving learner engagement and competence[5], a concept that has recently been proposed for medical education[4].

Due to the current healthcare workforce shortage across disciplines, there is an urgent need to expand training capacity. This challenge presents a fundamental chicken-and-egg dilemma as the need to train more medical students is confounded by an insufficient number of teaching physicians, creating a cyclical bottleneck in healthcare workforce development. Leveraging GenAI presents a promising solution to enhance clinical skills training without requiring proportional increases in faculty. However, realizing this potential

✉ Thomas Thesen
thomas.thesen@dartmouth.edu

1 Department of Medical Education, Geisel School of Medicine at Dartmouth, Hanover, USA

2 Department of Computer Science, Dartmouth College, Hanover, USA

3 Geisel School of Medicine at Dartmouth, MD Program, Hanover, USA

4 Research Computing, Dartmouth College, Hanover, USA

5 Department of Medicine, Geisel School of Medicine at Dartmouth, Hanover, USA

Published online: 12 August 2025

requires systematic testing in authentic educational environments and collecting and implementing learner feedback to develop effective AI-augmented pedagogical approaches.

A particularly promising application is the development of LLM-powered virtual patients that simulate clinical encounters with high conversational fidelity[6]. We created an AI Patient Actor application, an open-access web platform that enables students to practice communication skills and clinical reasoning in a controlled environment with immediate formative feedback[7]. This application, available open access at https://ai.dartmouth.edu/patient-actor, allows medical educators to use existing cases or create customized scenarios. This study examines the deployment of this tool among first-year medical students and analyzes their experiences and attitudes toward GenAI in clinical education.

We examine the use of LLMs within the context of three applicable educational theories, namely Situated Learning Theory, Deliberate Practice Theory, and Self-Determination Theory. *Situated Learning Theory* (SLT) posits that learning occurs most effectively when situated within authentic contexts and communities of practice [8]. The AI Patient Actor creates a simulated clinical environment that allows students to practice communication skills within a context that mirrors real clinical encounters. By engaging with virtual patients who present with realistic symptoms and respond conversationally, students can develop practical interviewing skills in an environment that replicates key aspects of clinical practice while removing the high-stakes pressure of real patient interactions. This balance provides a psychologically safe space for skill development, where mistakes become learning opportunities rather than potential patient harm.

According to *Deliberate Practice Theory* (DPT) optimal skill development requires focused, repetitive practice with immediate feedback targeted at specific aspects of performance[9]. This approach emphasizes the importance of engaging in structured activities designed to improve performance through gradual refinement rather than mere repetition. The AI Patient Actor application has the potential to support this framework by providing medical students with unlimited opportunities to practice clinical interviewing skills in a controlled environment that offers immediate formative feedback. Unlike traditional clinical experiences where feedback may be delayed or inconsistent, the GenAI system provides standardized assessment based on established rubrics, highlighting specific areas for improvement and enabling iterative refinement of communication techniques. This immediate feedback loop may allow students to identify gaps in their approach, adjust their interviewing strategies, and practice repeatedly until competence is achieved.

Self-Determination Theory proposes that optimal learning and engagement occur when three fundamental psychological needs are met: autonomy, competence, and relatedness[10]. The GenAI approach supports autonomy by allowing students to practice clinical interviewing at their own pace and convenience, without faculty scheduling constraints. The system has the potential to support competence development through structured practice opportunities with immediate feedback that helps students identify strengths and areas for improvement, gradually building clinical reasoning and communication skills. The combination of autonomy and competence —and confidence building— is likely to support intrinsic motivation, which would predict high engagement levels and positive attitudes expressed by students.

Methods

We enrolled first-year medical students ($n = 94$) from the "On Doctoring" course at Geisel School of Medicine at Dartmouth. Participants were required to engage with a minimum of three out of six available simulated patient cases over 12 weeks, resulting in 276 AI-mediated clinical encounters.

For the AI Patient Actor application, we utilized a Python-based web interface powered by GPT-4o that presents educator-created clinical cases while constraining LLM responses through carefully designed system prompts. The code is available below. The application processes student interactions through a Streamlit interface, with LangChain components optimizing information flow between the user, clinical case files, and the large language model. Voice capabilities were implemented using OpenAI's Whisper-3, allowing for both text and voice interactions while providing immediate formative feedback based on standardized rubrics adapted from our Objective Structured Clinical Exams (OSCE). Educator-created cases representing common disease presentations were used to avoid reliance on the LLM for clinical case generation and to ensure medical accuracy and educational relevance. Students conducted comprehensive interviews via text or voice interface, collecting complete medical histories at their convenience. Upon interview completion, they received physical examination findings and diagnostic results, developed differential diagnoses, and concluded with assessment and management plans as they would in actual clinical encounters. The system provided immediate formative feedback in written form based on standardized patient evaluation criteria (see Thesen et al. 2025[7] for more details).

The primary educational objective was for students to gather patient information conversationally and apply it to clinical note-writing that supported differential diagnosis development. Participants submitted three clinical notes at regular intervals for faculty evaluation. We administered surveys measuring students' comfort with GenAI in general

before and after the course. An additional end-of-course evaluation assessed perceived utility of GenAI in medical education and attitudes toward curriculum integration using 5-point Likert scales and open-ended questions. In previous cohorts, students completed similar assignments through in-person preceptor clinic encounters, but regularly experienced submission delays due to scheduling difficulties. These delays resulted in clustered submissions that limited opportunities for iterative feedback critical to clinical reasoning development[11, 12]. This challenge reflects nationwide difficulties in recruiting and retaining community preceptors[8,9]. The AI Patient Actor was implemented in Academic Year 2023–2024 to address these issues. We monitored submission timelines to evaluate improvements in timeliness and gathered data on student perspectives, recognizing that current learners will influence future GenAI integration in clinical practice[13]. The study was categorized as exempt by the Dartmouth College IRB.

Analysis

We collected quantitative and qualitative survey responses from all participants ($n = 94$) through Likert-scale questionnaires and open-ended text fields.

Sentiment Analysis

Student text comments underwent sentiment analysis using the TextBlob Python library, which calculated polarity scores ranging from − 1.0 (maximally negative) to + 1.0 (maximally positive) for each response[14]. These scores were derived using lexicon-based sentiment dictionaries and linguistic rules that evaluate textual positivity, negativity, or neutrality. The algorithm weighted positive terms (e.g., "helpful") to contribute favorably to sentiment scores, while negative terms (e.g., "frustrating") diminished scores.

Content Analysis

Our content analysis methodology employed a two-stage human-AI collaborative approach[15, 16]. First, we utilized ChatGPT-4o via API interface to identify preliminary thematic patterns across student feedback. Subsequently, two researchers systematically verified all AI-identified concepts against original data, refining concept definitions and supplementing categories missed by the model. This process yielded six distinct concept categories: (1) opportunities for practice and reflection, (2) effective immediate feedback, (3) technical issues and frustrations, (4) enhanced clinical reasoning and diagnostic skills, (5) utility in practicing clinical write-ups, and (6) limitations in simulating real interactions. We then preprocessed student comments into discrete sentences, with each independently classified by GPT-4 into one of these six predefined categories. Classifications were conducted via the API with temperature set to zero to ensure consistent categorization. Sentences lacking clear category alignment were labeled "Uncategorized." All automated classifications underwent manual review by the researchers.

Results

Survey

Following AI Patient Actor implementation, we observed a 75.9% reduction in late assignment submissions compared to previous cohorts. Among the 69 students who completed evaluations, 61% reported positive experiences: 3% rated the AI cases as "Good," 30% as "Very Good," and 14% as "Excellent," while 14% considered them "Fair" and 7% "Poor." Students' mean comfort level with GenAI in general increased significantly from 61.0% (SD = 24.4) pre-intervention to 76.0% (SD = 11.6) post-intervention ($t = -3.74$, $p < 0.001$), with perceived utility of AI averaging 73.9% (SD = 20.1). Participants generally endorsed artificial intelligence as a valuable tool for preclinical communication skills practice, with a mean agreement score of 3.27 (SD = 1.37) on a 5-point scale (see Fig. 1A). The combination of increased general comfort with GenAI and high perceived utility scores reflects favorable student attitudes toward the GenAI educational technology.

Sentiment Analysis

Sentiment analysis of student responses showed a broad spectrum of reactions, with polarity scores ranging from − 0.25 (negative) to 0.75 (positive). The distribution was predominantly positive, with fewer negative sentiments (see Fig. 1B). Positive feedback emphasized four key benefits of the AI Patient interactions: enhanced clinical reasoning capabilities, flexible practice opportunities unrestricted by scheduling constraints, improved clinical documentation skills, and valuable immediate feedback. Negative comments centered on four primary concerns: limited transferability to real-world clinical skills, insufficient realism in patient interactions, occasional inaccuracies in automated feedback, and the inherently impersonal nature of AI-based patient encounters.

Content Analysis

Our content analysis identified six major concepts within student feedback, distributed across positive and negative sentiments (Fig. 1C). Positive themes, ranked by frequency, included opportunities for practice and reflection, clinical

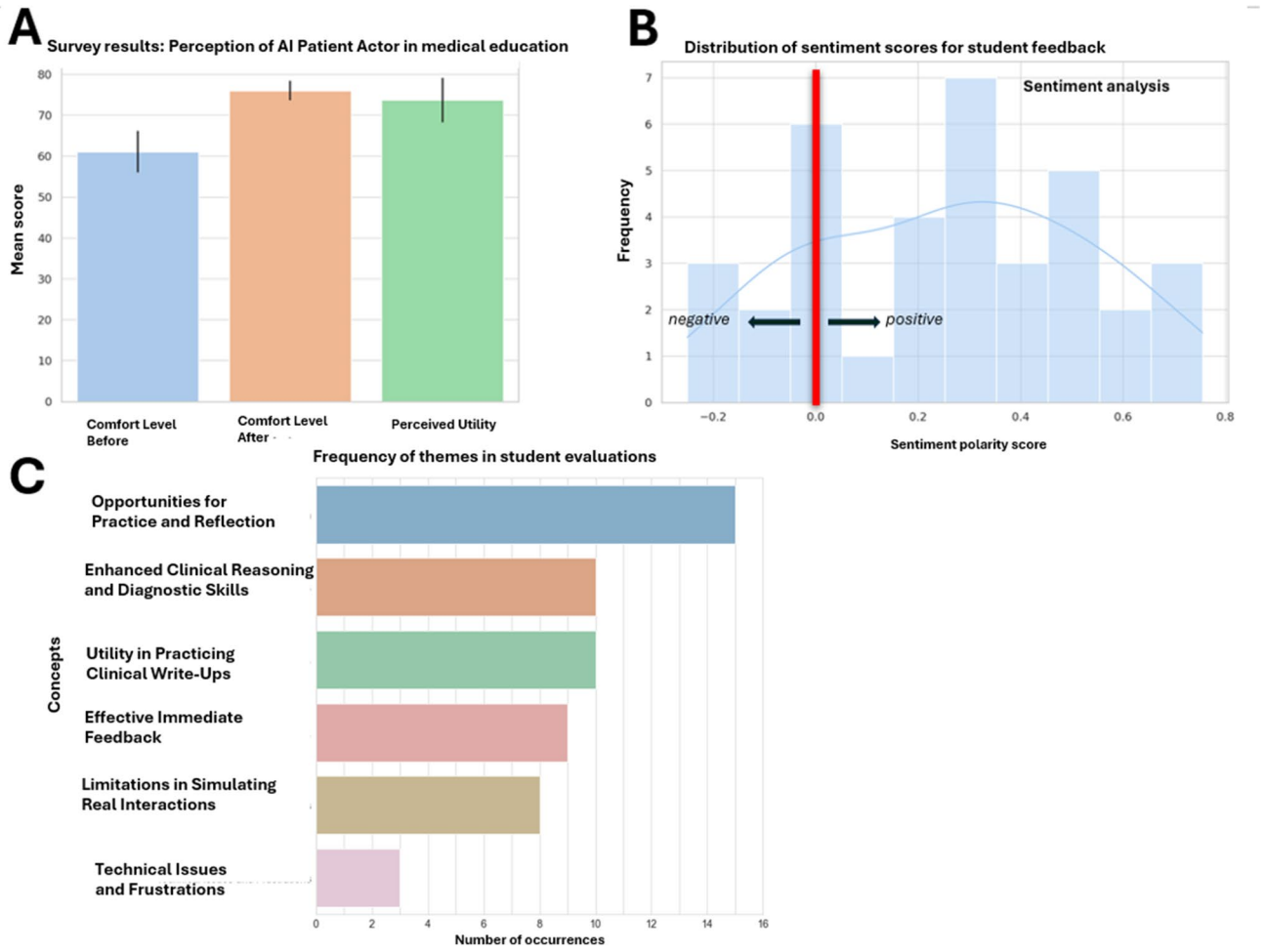

Fig. 1 **A** Mean scores with standard error bars for survey questions. **B** Distribution of sentiment polarity scores for written student feedback. Sentiment polarity scores range from − 1 (most negative) to + 1 (most positive). The histogram shows the frequency of responses across different sentiment scores, while the density curve (KDE) indicates the overall spread of sentiment. **C** Frequency of major concepts in student feedback

reasoning and diagnostic skill development, utility for practicing clinical documentation, and effective immediate feedback. Negative themes comprised limitations in simulating authentic patient encounters and technical issues with the application interface. Representative student comments for each theme are presented in the Appendix. (For details on the analysis libraries and parameters, please refer to the analysis code available in the Code & Data Availability section below.)

Discussion

Medical students demonstrated predominantly positive attitudes toward the AI Patient Actor, highlighting its benefits for flexible practice, immediate feedback, and clinical communication skills training. They valued the low-pressure environment that enabled self-paced learning, particularly beneficial for those requiring additional time to develop confidence before real patient encounters. The immediate feedback facilitated rapid identification of improvement areas, reinforcing effective behaviors and supporting iterative learning. Additionally, the GenAI implementation significantly reduced late submissions by eliminating scheduling constraints and increasing learner autonomy. These findings suggest AI Patient Actors may effectively bridge the gap between theoretical knowledge and practical clinical skill development in early medical education.

Nevertheless, some students expressed reservations regarding the limited authenticity of AI encounters compared to human interactions and questioned their effectiveness in building comprehensive clinical skills. These mixed perspectives underscore the importance for medical educators to design developmentally appropriate

exercises and clearly communicate both capabilities and limitations of GenAI applications while establishing realistic expectations. Findings show that incorporating AI Patient Actors into early medical education significantly improves students' comfort with GenAI technology overall and enhances their perception of its educational value. The structured exposure provided through these interactions effectively reduces initial hesitation or uncertainty with the new technology and may simultaneously build students' confidence in applying AI within clinical contexts.

Students broadly endorsed the AI Patient Actor as valuable for preclinical communication skills development, aligning with previous research on AI-simulated patient encounters and personalized feedback[17]. Given widespread challenges in recruiting clinical preceptors, GenAI presents an opportunity to address specific curricular objectives. Student enthusiasm, combined with documented gaps in understanding AI's clinical applications, highlights the need for comprehensive AI education within medical curricula.

The high perceived utility for GenAI supports its potential benefits for clinical decision-making and communication skills training[18]. However, variability in comfort and utility ratings indicates differing acceptance levels among students, potentially influenced by prior technology exposure, AI familiarity, or learning preferences, factors that should inform future GenAI tool development.

Our results can be explained in the context of the three learning theories introduced earlier. When we consider *Situated Learning Theory*, students responded so positively because they were practicing in an environment that felt, to a degree, like a real clinical setting but without the pressure of potentially harming an actual patient. It gave them that safe space to try out their skills within the right context. Considering *Deliberate Practice Theory*, we believe the immediate feedback was most important. Students specifically mentioned how helpful it was to get instant guidance on what they did well and what needed work, which let them improve their interviewing techniques without waiting for potentially delayed faculty feedback.

In alignment with *Self-Determination Theory*, students felt that the LLM approach allowed for students to have autonomy, develop competence, and, to some degree, experience relatedness because students could practice whenever they wanted, they could see themselves improving, and they still got to engage in professional-type conversations. The fact that students noticed limitations in how "real" the AI felt shows that while AI helps build foundational skills, human interaction remains essential in medical training.

Limitations include our relatively small single-institution sample and reliance on self-reported measures. Future investigations should employ larger, more diverse cohorts with longitudinal designs to track evolving attitudes toward GenAI and directly measure clinical performance outcomes from AI-supported training.

In conclusion, our findings present a promising perspective on GenAI's potential to help medical trainees develop essential clinical skills. Structured and evolving interactions with these technologies early in training may help demystify AI, increase students' willingness to engage with it clinically, and better prepare them for healthcare environments that are increasingly incorporating AI-augmented solutions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40670-025-02473-x.

Funding Supported by the American Medical Association Change-MedEd Initiative for Precision Education and the American Academy of Neurology.

Code and Data Availability The code for the AI patient Actor app and data analysis is publicly available at https://github.com/tominny/ai-patient-actor-ondoc. The deployed app is available open-access at https://ai.dartmouth.edu/patient-actor.

Declarations

Competing interests The authors declare no competing interests.

References

1. Mir MM, Mir GM, Raina NT, et al. Application of artificial intelligence in medical education: current scenario and future perspectives. J Adv Med Educ Prof. 2023;11(3):133–40. https://doi.org/10.30476/JAMP.2023.98655.1803.
2. Feng S, Shen Y. ChatGPT and the future of medical education. Acad Med. 2023;98(8):867. https://doi.org/10.1097/ACM.0000000000005242.
3. Park SH, Pinto-Powell R, Thesen T, et al. Preparing healthcare leaders of the digital age with an integrative artificial intelligence curriculum: a pilot study. Med Educ Online. 2024;29(1): 2315684. https://doi.org/10.1080/10872981.2024.2315684.
4. Triola MM, Burk-Rafel J. Precision medical education. Acad Med J Assoc Am Med Coll. 2023;98(7):775–81. https://doi.org/10.1097/ACM.0000000000005227.
5. Hart SA. Precision education initiative: moving toward personalized education. Mind Brain Educ. 2016;10(4):209–11.
6. Holderried F, Stegemann-Philipps C, Herrmann-Werner A, et al. A language model–powered simulated patient with automated feedback for history taking: prospective study. JMIR Med Educ. 2024;10: e59213. https://doi.org/10.2196/59213.
7. Thesen T, Alilonu N, Stone S. AI Patient Actor: an open-access Generative-AI app for communication training in health professions. Med Sci Educ 2025;. Published online [In press].
8. Lave J, Wenger E. Situated learning: Legitimate peripheral participation. Cambridge: Cambridge University Press; 1991.
9. Krampe RT, Ericsson KA. Maintaining excellence: deliberate practice and elite performance in young and older pianists. J Exp Psychol Gen. 1996;125(4):331.
10. Deci EL, Ryan RM. Intrinsic motivation and self-determination in human behavior. Berlin: Springer Science & Business Media. 1985. https://doi.org/10.1007/978-1-4899-2271-7.
11. Versteeg M, Hendriks RA, Thomas A, Ommering BWC, Steendijk P. Conceptualising spaced learning in health professions

education: a scoping review. Med Educ. 2020;54(3):205–16. https://doi.org/10.1111/medu.14025.
12. Nieman LZ, Cheng L, Hormann M, Farnie MA, Molony DA, Butler P. The impact of preclinical preceptorships on learning the fundamentals of clinical medicine and physical diagnosis skills. Acad Med J Assoc Am Med Coll. 2006;81(4):342–6. https://doi.org/10.1097/00001888-200604000-00007.
13. Liu DS, Sawyer J, Luna A, et al. Perceptions of US medical students on artificial intelligence in medicine: mixed methods survey study. JMIR Med Educ. 2022;8(4): e38325. https://doi.org/10.2196/38325.
14. TextBlob: Simplified Text Processing — TextBlob 0.18.0.post0 documentation. October 29, 2024. https://textblob.readthedocs.io/en/dev/. Accessed 28 Oct 2024
15. Prescott MR, Yeager S, Ham L, et al. Comparing the efficacy and efficiency of human and generative ai: qualitative thematic analyses. JMIR AI. 2024;3(1): e54482. https://doi.org/10.2196/54482.
16. Nguyen-Trung K, Nguyen N. Narrative-Integrated Thematic Analysis (NITA): AI-supported theme generation without coding. Preprint. 2025. https://doi.org/10.2139/ssrn.5163884.
17. Hamilton A. Artificial intelligence and healthcare simulation: the shifting landscape of medical education. Cureus. 2024;16(5): e59747. https://doi.org/10.7759/cureus.59747.
18. Chen M, Zhang B, Cai Z, et al. Acceptance of clinical artificial intelligence among physicians and medical students: a systematic review with cross-sectional survey. Front Med. 2022;9: 990604.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.