Graduate Medical Education Research Journal

Manuscript 1640

AI-Based Simulated Patient Encounters for Clinical Reasoning: A Needs Assessment

Alex Sorrick

David Liebovitz MD

Sarah Vick MD

Andrew S. Parsons MD MPH

Brian W. Locke MD MSc

See next page for additional authors

AI-Based Simulated Patient Encounters for Clinical Reasoning: A Needs Assessment

Abstract

Background and Objectives

Artificial intelligence offers an opportunity to enhance formative assessment of clinical reasoning skills. Traditionally, these skills have been assessed through workplace-based assessment or simulated patient encounters, both of which are limited by variability in case exposure, faculty resources, scheduling logistics, and cost. AI-based simulated patient encounters provide a scalable solution for convenient, low-stakes practice. At the University of Nebraska Medical Center (UNMC), we piloted an AI-based simulated patient case and conducted a needs assessment to explore current practices and user experiences.

Methods

From June to August 2025, participants trialed an AI-based simulated patient case and received feedback on their clinical reasoning. They then completed a brief 20-question survey assessing acceptability of the technology, perceived educational value, potential future applications, and comparison to existing assessment modalities. Invitations to participate were sent to medical students (M1-M4), residents (PGY1-PGY3), and attending physicians in internal medicine, family medicine, and emergency medicine at UNMC.

Results

Thirty-four surveys were completed: 29% pre-clinical students (M1-M2), 32% clinical students (M3-M4), 12% residents, and 26% attending physicians. Most respondents found the experience acceptable. In particular, the AI-based clinical reasoning feedback was perceived positively, with most agreeing it identified gaps, provided actionable steps, and was useful for learning. While opinions on replacing traditional Observed Structured Clinical Examinations (OSCEs) with AI-based encounters were mixed, the majority expressed interest in using this technology to improve their clinical reasoning skills.

Conclusions

AI-based simulated patient encounters offer a feasible way to supplement and scale current methods in clinical reasoning assessment. While not intended to replace human encounters, they provide additional opportunities for on-demand, low-pressure practice – an especially valuable tool for busy trainees. Rather than viewing AI as a threat to clinical reasoning, personalized tools should be implemented and studied to strengthen, advance, and guide its development.

Keywords

clinical reasoning, artificial intelligence, patient simulation, medical education, formative assessment, needs assessment

Authors

Alex Sorrick, David Liebovitz MD, Sarah Vick MD, Andrew S. Parsons MD MPH, Brian W. Locke MD MSc,

Jonathan Herskovitz MD PhD, and Cory Rohlfsen MD

Title: AI-Based Simulated Patient Encounters for Clinical Reasoning: A Needs Assessment

Shortened title (40 characters): AI-Based Simulated Encounters for Clinical Reasoning Assessment

Authors:

Alex Sorrick MD, University of Nebraska Medical Center

David Liebovitz MD, Northwestern University

Sarah Vick MD, University of Kentucky

Andrew S Parsons MD MPH, University of Virginia

Brian W Locke MD MSc, Intermountain Health

Jonathan Herskovitz MD PhD, University of Nebraska Medical Center

Cory Rohlfsen MD, University of Nebraska Medical Center

The authors have no conflicts of interest to report.
Funding: This work was supported in part by an Educational Support Grant through UNMC College of Medicine.

Ethics statement: This study was reviewed by the Institutional Review Board (IRB) and determined to be exempt (IRB protocol 0269-25-EX). Participation was voluntary and informed consent was obtained. Data was de-identified prior to analysis to protect confidentiality.

Data availability statement: All data supporting the findings of this study are reported within the manuscript and supplemental file. Individual-level transcript data are not available due to participant confidentiality.

Author contributions: CR conceived and designed the study. DL developed the AI-based simulation and analytic framework. JH and AS oversaw data collection. SV, AP, BL, AS, JH, CR performed quantitative and qualitative analyses. AS drafted the manuscript. All authors critically revised the manuscript and approved the final version.

Abstract

Background and Objectives:

Artificial intelligence offers an opportunity to enhance formative assessment of clinical reasoning skills. Traditionally, these skills have been assessed through workplace-based assessment or simulated patient encounters, both of which are limited by variability in case exposure, faculty resources, scheduling logistics, and cost. AI-based simulated patient encounters provide a scalable solution for convenient, low-stakes practice. At the University of Nebraska Medical Center (UNMC), we piloted an AI-based simulated patient case and conducted a needs assessment to explore current practices and user experiences.

Methods:

From June to August 2025, participants trialed an AI-based simulated patient case and received feedback on their clinical reasoning. They then completed a brief 20-question survey assessing acceptability of the technology, perceived educational value, potential future applications, and comparison to existing assessment modalities. Invitations to participate were sent to medical students (M1-M4), residents (PGY1-PGY3), and attending physicians in internal medicine, family medicine, and emergency medicine at UNMC.

Results

Thirty-four surveys were completed: 29% pre-clinical students (M1-M2), 32% clinical students (M3-M4), 12% residents, and 26% attending physicians. Most respondents found the experience acceptable. In particular, the AI-based clinical reasoning feedback was perceived positively, with most agreeing it identified gaps, provided actionable steps, and was useful for learning. While opinions on replacing traditional Observed Structured Clinical Examinations (OSCEs) with AI-based encounters were mixed, the majority expressed interest in using this technology to improve their clinical reasoning skills.

Conclusions:

AI-based simulated patient encounters offer a feasible way to supplement and scale current methods in clinical reasoning assessment. While not intended to replace human encounters, they provide additional opportunities for on-demand, low-pressure practice – an especially valuable tool for busy trainees. Rather than viewing AI as a threat to clinical reasoning, personalized tools should be implemented and studied to strengthen, advance, and guide its development.

Key words: clinical reasoning, artificial intelligence, patient simulation, medical education, formative assessment, needs assessment

Introduction:

Artificial intelligence (AI) presents a unique opportunity to provide low-stakes practice for developing clinical reasoning skills. Traditional assessment methods like workplace-based assessments (WBAs) and Observed Structured Clinical Examinations (OSCEs) are valuable but limited by case variability, faculty resources, and administrative costs.[1,2] In one study of 180 residents across seventeen specialties, only 25% reported receiving regular feedback on their performance.[3] Asynchronous simulated patient encounters may broaden case exposure and increase feedback opportunities, but little is known about trainees' perceptions of unsupervised, AI-generated feedback. While summative OSCEs are the gold standard for assessing low-frequency, high-risk events (e.g., resuscitation) or entrusting baseline readiness skills for incoming interns (e.g. patient communication), feedback on *clinical reasoning* is often delayed and generalized from a human-graded rubric rather than customized in real time and at scale.[4,5] AI-simulated encounters with interactive "ChatBot as patient" personas could bridge this gap by automating *formative* assessment of more situated skills like information gathering and developing a differential diagnosis. A lack of expert-appraised, self-directed, AI learning platforms tailored for on-demand practice of these critical skills prompted this institutional-specific, educational needs assessment.

Methods:

From June to August 2025, participants at the University of Nebraska Medical Center trialed an AI-based simulated patient case and received immediate feedback on their clinical reasoning. They then completed a brief 20-question survey. Participants included medical students (M1-M4), residents (PGY1-PGY3), and attending physicians in internal medicine, family medicine, and emergency medicine. Survey items included a 4-point Likert scale (strongly agree, agree, disagree, strongly disagree) and open-ended responses.

Results:

Thirty-four surveys were completed across a spectrum of demographics including 29% pre-clinical students (M1-M2), 32% clinical students (M3-M4), 12% residents (PGY1-3), and 26% attending physicians, with a majority (85%) from internal medicine (table 1). Most participants used text responses (74%) while 26% used audio or both.

Overall, nearly one quarter (23.56%) reported not having previously received feedback on their clinical reasoning skills. Most participants agreed that the AI-based interface was easy to understand (85.3%), authentic (70.6%), and a satisfying experience (94.1%). A majority of participants also agreed that the app identified areas for improvement (94.2%), provided clear action steps (85.3%) and was useful for learning (97%). Despite varied training levels, most participants agreed that the app was appropriate for their level of training (88.3%) and felt confident in applying what they learned to real patient care (85.3%) (table 2).

For those who felt the app could be more authentic, respondents primarily cited the brevity of responses and lack of non-verbal cues as limiting the conversational flow. One participant stated, "It felt like pulling teeth at times, whereas a spontaneous conversation with a human allows for more natural exploration of their disease" (supplemental file 1).

Opinions on using this technology compared to traditional summative OSCEs with standardized patients were mixed. Supportive responses noted the app experience was less anxiety-provoking and more standardized compared to traditional OSCEs. One medical student responded, "This would be great to do as an OSCE because it would consistently grade people and grading would be done super fast. The feedback was really good and I think it graded my performance more than fairly." Opposing responses emphasized the importance of the human aspect. One resident participant responded, "This is great practice for clinical reasoning and learning, but not for the soft skills - working with people" (supplemental file 1).

Notably, only 56% of participants solved the case and even those with the most clinical experience commented on the challenging nature of the experience. One attending physician stated, "The case presented would never be solved within one encounter and felt too specific." Despite the challenging nature of the case, 88.3% of participants agreed that they would use this for deliberate practice in the future. One attending participant highlighted the skill of adaptability: "I think the technology is fairly remarkable [...] could be a helpful means of giving clinical reps to learners." A trainee participant also wrote, "It was nice to practice my clinical reasoning in a non-timed/less stressful experience. I was able to think through everything before taking next steps" (supplemental file 1).

Discussion:

Our needs assessment highlights the potential for integrating AI-based tools into clinical reasoning assessment and feedback. While AI-augmented feedback is unsupervised, our pilot innovation shows educational promise for participants across multiple training levels to practice their clinical reasoning skills within a curated LLM-simulated patient encounter, likely owing to the immediate, personalized, and actionable feedback received.

We programmed the simulated patient with a challenging case and terse persona to test the limits of more skilled diagnosticians. The intention was not to improve bedside clinical skills or communication. Despite this narrow focus and the limited "diagnostic success" of participants, it's notable that favorability in the experience was consistent across participant levels. Individualized feedback with actionable steps for real-world application speaks to the potential of AI-based simulation experiences to bridge major gaps in precision learning. Importantly, most respondents expressed interest in using this technology for deliberate practice.

Given the primacy of OSCEs to assess communication skills with standardized patients, it's not surprising opinions were mixed as to whether there is a role for AI-based, summative assessment. The tradeoff seems to be fidelity of human interaction (non-verbal cues, authenticity,

flow) and standardization of AI outputs (perceived as less biased yet robotic). A more formal qualitative "faithfulness" evaluation of OSCE versus AI transcripts would be needed to explore this important question of *standardized responses* although human-ness was not our stated goal. We suspect the challenging case may also have skewed participant perception of utility in summative assessment as rates of diagnostic success are much higher in traditional OSCEs.

Limitations include low response rate, single-center design, and internal medicine predominance, which may limit generalizability to other programs and institutions.

Conclusion:

While not intended to replace human encounters, AI-based patient simulations provide low-stakes, on-demand practice for trainees to develop their clinical reasoning skills at scale. As patient care grows increasingly complex, embracing innovative technologies to train the next generation of physicians will be an important step towards self-directed, precision learning.

References:

[1] Daniel M, Rencic J, Durning S, et al. Clinical Reasoning Assessment Methods: A Scoping Review and Practical Guidance. Academic Medicine 94(6):p 902-912, June 2019.

[2] Elendu C, Amaechi D, Okatta A, et al. The impact of simulation-based training in medical education: A review. Medicine 103(27):p e38813, July 2024.

[3] Shafian S, Ilaghi M, Shahsavani Y, et al. The feedback dilemma in medical education: insights from medical residents' perspectives. BMC Med Educ. April 2024.

[4] Mangold KA, Jeffers JM, Burns RA, et al. An Objective Structured Clinical Examination to Improve Formative Assessment for Senior Pediatrics Residents. J Grad Med Educ. September 2015.

[5] CarlLee S, Rowat J, Suneja M. Assessing Entrustable Professional Activities Using an Orientation OSCE: Identifying the Gaps. J Grad Med Educ. April 2019.