

Introduction

Effective and empathetic communication is central to patient-centred care and is necessary for building a trusting patient-physician relationship.¹ Therefore, the effective development of communication skills during medical school and residency is essential to becoming a well-rounded physician.¹

Communication skills of internal medicine (IM) residents are assessed during the Objective Structured Clinical Examinations (OSCEs) that involve stimulating clinical interactions with standardized patients (SPs).² However, coordinating SPs and physician assessors is costly and time-consuming, limiting training opportunities for IM residents.³

Generative artificial Intelligence (AI) has the potential to improve medical education by giving learners more opportunities to practice standardized patient encounters to develop their communication skills.

Objectives

The primary objectives were to:

- 1) To compare AI versus physician assessments of resident performance in an IM OSCE station, and
- 2) To survey resident and physician evaluator perspectives on the use of AI in medical education.

Methods

IM residents participated in a formative OSCE station with a commercially available AI-powered telephone chatbot and assessment tool developed by Osler AI, Inc.

OSCE Case	The AI-powered OSCE station was a clinical scenario on the disclosure of a medical error to an AI substitute decision maker.
OSCE Scoring Rubric	Scoring rubrics included a 7-item checklist (CL), a 6-item communication skill scoring scale, and a global rating scale (GRS).
Survey Design	A cross-sectional survey investigated the perceptions of IM residents and physicians regarding AI in medical education following the AI-simulated OSCE case.
Data Analysis	AI vs physician scores were compared using mean differences. Borderline regression generated passing scores for each assessment group.

Results

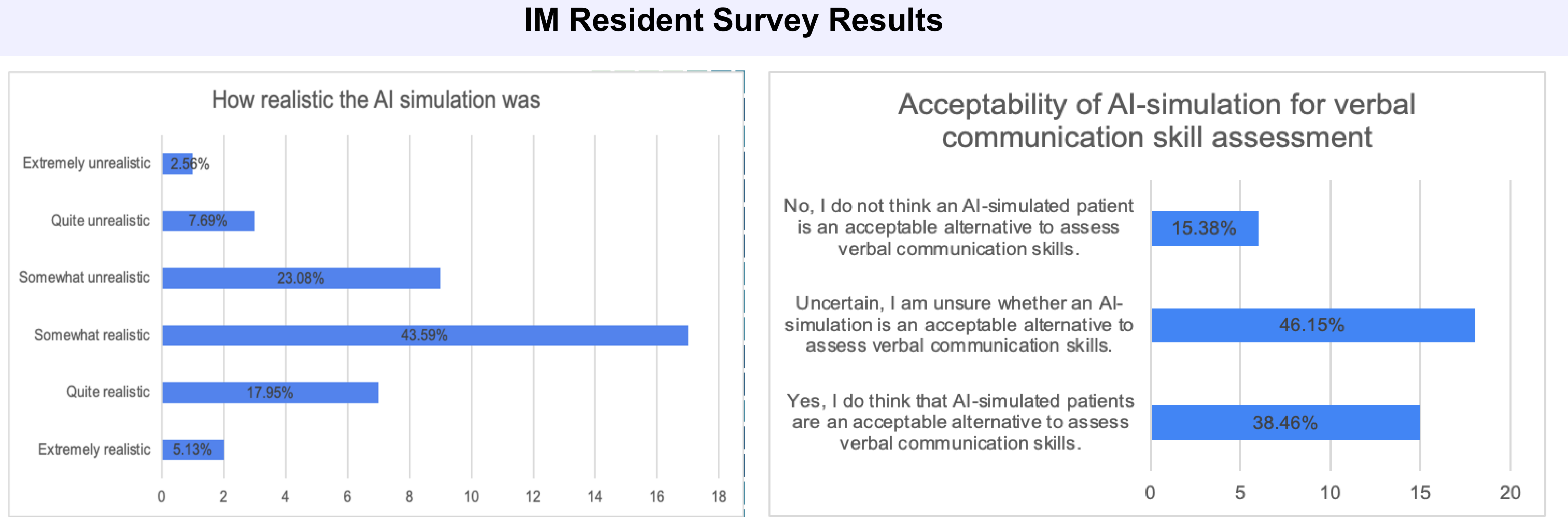


Figure 1. IM resident opinions on the realism of the AI-based OSCE station

Figure 2. IM resident opinions on AI use in OSCE assessments

Physician Survey Results

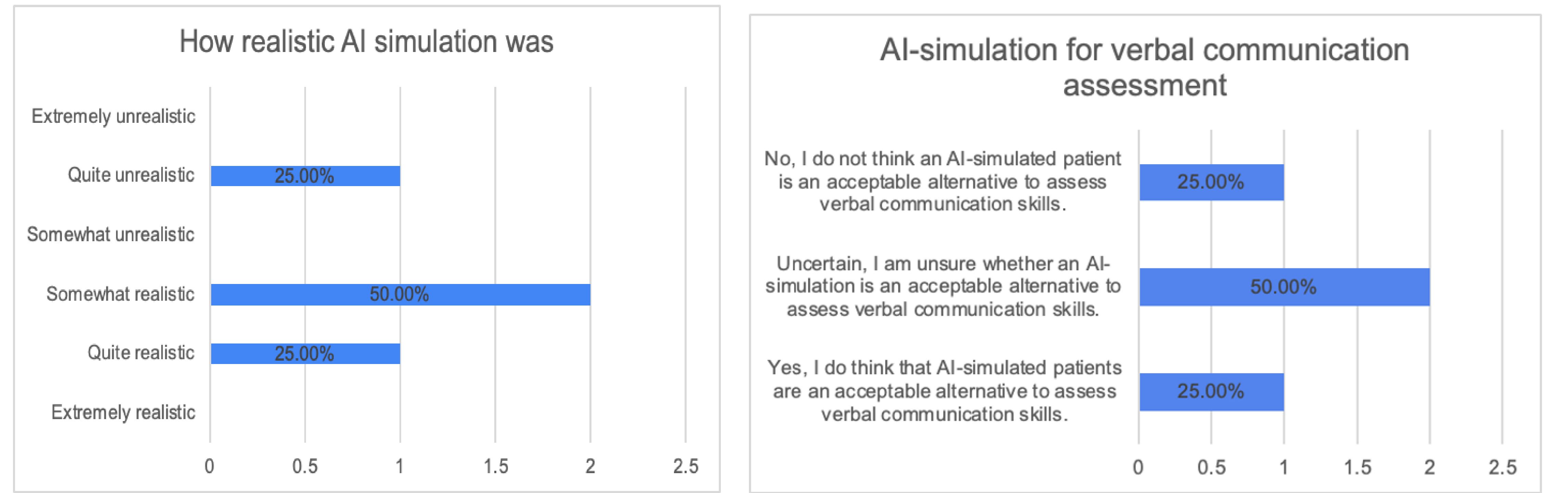


Figure 3. Physician opinions on the realism of the AI-based OSCE station

Figure 4. Physician openness to AI use in different areas of medicine overall

AI vs Physician OSCE Scores

Communication Skills				
Category	Mean Difference (Physician – AI)	Score Direction	P-value	Significance (p<.05)
Listening Skills	-0.057	AI Higher	0.646	No
Questioning Skills	0.200	AI Lower	0.187	No
Rapport with Patient	0	No Difference	1.000	No
Information Giving	-0.314	AI Higher	0.157	No
Professional Behaviour	-0.400	AI Higher	0.005	Yes
Ethical Conduct	0.143	AI Lower	0.131	No

Figure 5. Physician compared to AI assessment of IM resident OSCE communication skills

Results

AI vs Physician OSCE Scores				
Content Checklist				
Category	Mean Difference (Physician – AI)	Score Direction	P-value	Significance (p<.05)
C1	0.343	AI Lower	0.002	Yes
C2	-0.057	AI Higher	0.311	No
C3	-0.029	AI Higher	0.792	No
C4	-0.086	AI Higher	0.083	No
C5	-0.114	AI Higher	0.261	No
C6	0.029	AI Lower	0.648	No
C7	0.029	AI Lower	0.812	No

Figure 6. Physician compared to AI assessment of IM resident OSCE content performance

Overall Ratings				
Category	Mean Difference (Physician – AI)	Score Direction	P-value	Significance (p<.05)
Global performance rating	-0.371	AI Higher	0.064	No
Overall score	0.000	AI Lower	1.000	No

Figure 7. Physician compared to AI assessment of IM resident OSCE overall performance

Conclusion

It is feasible to use an AI-based clinical simulation both as an SP and for learner assessment in IM OSCEs.

Our encouraging results demonstrate that generative AI and physician assessors give IM residents similar OSCE scores in most categories. More research is required to ensure the validity and reliability of AI-based clinical simulation and assessment tools prior to their implementation in communication skills training.

Acknowledgements

This study was funded by The Ottawa Hospital Department of Medicine and supported by the Ottawa Hospital Research Institute. In-kind support was generously provided by OslerAI.

References

1. Hojat M, Louis DZ, Maxwell K, Markham F, Wender R, Gonnella JS. Patient perceptions of physician empathy, satisfaction with physician, interpersonal trust, and compliance. *Int J Med Educ.* 2010;1:83-87. doi:10.5116/ijme.4000.5701.

2. Gilligan C, Powell M, Lynagh MC, et al. Interventions for improving medical students' interpersonal communication in medical consultations. *Cochrane Database of Systematic Reviews.* 2021;2021(2). doi:10.1002/14651858.CD012418.pub2

3. Nestel D, Tabak D, Tierney T, et al. Key challenges in simulated patient programs: An international comparative case study. *BMC Med Educ.* 2011;11(1):69. doi:10.1186/1472-6920-11-69

Evidence of Conflict of Financial Interest

	Co-author	Conflict disclosures
1	Kyran Sachdeva	None
2	Samantha Halman	None
3	Brett Hryciw	Dr. Hryciw is a co-founder of OslerAI Inc. Dr Hryciw was not involved in handling any of the raw data prior to analysis. Dr Hryciw and Osler AI did not receive any financial compensation for this study. Osler AI did not use any data from this study for the purposes of product development or marketing.
4	Kordan Harvey	Dr. Harvey s a co-founder of OslerAI Inc. Dr Harvey was not involved in handling any of the raw data prior to analysis. Dr Harvey and Osler AI did not receive any financial compensation for this study. Osler AI did not use any data from this study for the purposes of product development or marketing.
5	Nicole Hryciw	Dr. Hryciw is a co-founder of OslerAI Inc. Dr Hryciw was not involved in handling any of the raw data prior to analysis. Dr Hryciw and Osler AI did not receive any financial compensation for this study. Osler AI did not use any data from this study for the purposes of product development or marketing.