Research Letter

Self-Directed Use of General-Purpose AI Tools for History-Taking Practice Among Clinical-Phase Medical Students: Cross-Sectional Survey

Krishan Goyal*, MBBS, BSc (Hons), PGCert (ClinEd); Shahmeer Noori*, BMBS, iBSc, PGCert (MedEd); Sandiso Moyo, BMBS, MSc, MRCGP, PGCert (ClinEd); Neeraj Kumar Malhan, MBBS, BSc (Hons), FRCEM, PGCert (ClinEd)

Imperial College School of Medicine, Imperial College London, London, England, United Kingdom
*these authors contributed equally

Corresponding Author:

Krishan Goyal, MBBS, BSc (Hons), PGCert (ClinEd)
Imperial College School of Medicine
Imperial College London
Reynolds Building, St Dunstan's Road, Charing Cross Campus
London, England W6 8RP
United Kingdom
Phone: 44 02075947259
Email: kg217@ic.ac.uk

Abstract

A cross-sectional survey of 25 year 3 medical students at a UK medical school showed that most students had experimented with general-purpose AI for objective structured clinical examination history-taking practice, but few used it regularly. Students ranked general-purpose AI last among other revision resources, citing poor realism and unreliable feedback.

JMIR Form Res 2026;10:e108381; doi: 10.2196/108381

Keywords: artificial intelligence; medical education; ChatGPT; large language models; medical students; history taking; clinical skills; simulated patient; self-directed learning; cross-sectional survey; chatbot; conversational agent

Introduction

History taking is a core clinical skill assessed among medical students in objective structured clinical examinations (OSCEs) [1,2]. Student practice is limited by timetabling, actor availability, and clinical demands [3]. Purpose-built large language models (LLMs) can plausibly simulate patients [4], provide feedback aligning with human raters [5], and improve students' history-taking scores [6]. However, there is limited literature on how students use generic LLMs (eg, ChatGPT) for this purpose. This is an important gap to understand, given most health care students worldwide receive no formal AI education [7]. Hence, we surveyed year 3 medical students on the prevalence, patterns, and perceived value of self-directed AI use for OSCE history-taking practice.

Methods

Study Design and Analyses

We conducted a cross-sectional, mixed methods survey using nonprobability convenience sampling. All year 3 Imperial College London medical students with clinical placements at Hillingdon Hospital between January and June 2026 were invited to participate (n=37). Participants were invited via brief announcements after teaching sessions, with a QR code linking to an anonymous Microsoft Forms questionnaire (Multimedia Appendix 1). Participation was welcomed irrespective of AI use, and the survey was open from May 26 to June 10, 2026, closing 3 weeks prior to their OSCEs. Reporting followed the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist (Checklist 1) [8].

The questionnaire, informed by prior AI-simulated patient studies [4,5], covered 3 domains: general AI use (platforms, duration, and training); OSCE-preparation use (frequency, comparison with traditional revision, and ranking of 5

resources); and history-taking–specific use (activities and Likert-scale realism/confidence), plus 1 free-text item on limitations. Although not formally piloted, the questionnaire was iteratively refined through discussion and feedback with the research team and clinical teaching fellow colleagues.

Quantitative data were analyzed descriptively (frequencies/percentages or mean ranks for ranking items) without inferential testing. Free-text responses underwent inductive thematic analysis [9], with codes generated independently by 2 researchers and discrepancies resolved by discussion.

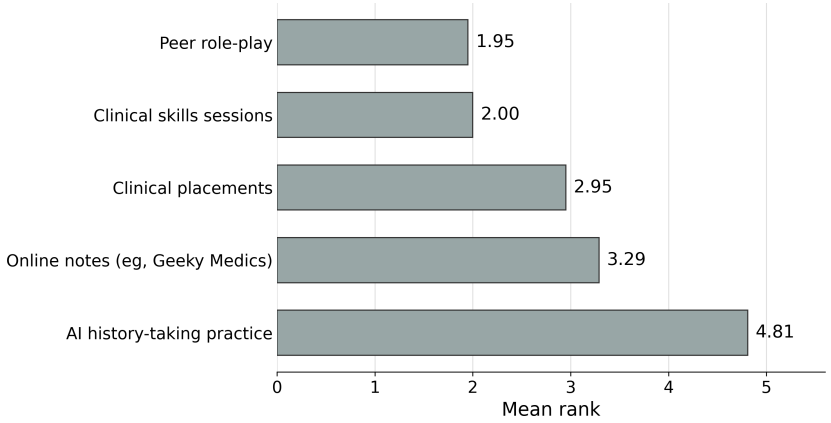
Ethical Considerations

The Imperial College Education Ethics Review Process approved the study (EERP2526-111); an information sheet and consent statement preceded the questionnaire, and no identifying or demographic data were collected.

Results

Of 37 eligible students, 25 responded (67.6%). Of the 25 respondents, 22 (88%) used AI for their studies, yet only 2 (8%) had received formal AI teaching. The most popular AI platforms included ChatGPT (21/22), Claude (8/22), and Gemini (4/22); multiple selections were permitted. Of the 25 respondents, 21 (84%) had tried or used AI for history-taking practice; however, 13 of 21 (62%) had tried it and stopped and only 2 of 21 (10%) used it regularly. Simulated patient role-play was the most common application (11/21), followed by OSCE scenario generation (9/21) and symptom/red-flag learning (9/21); using AI for communication skills practice, such as breaking bad news, was rare (2/21). Students ranked AI last among 5 popular alternative OSCE revision resources (Figure 1). No student rated AI as more useful than traditional methods; of 21 students, 10 (48%) rated it "much less useful" and 6 (29%) "less useful."

Figure 1. Mean student ranking of 5 OSCE revision resources for history-taking practice (1=most useful; 5=least useful).



Perceived realism was evenly split (agree: 7/21; neutral: 7/21; disagree: 7/21); perceived benefit to OSCE confidence skewed positive (agree: 9/21; neutral: 6/21; disagree: 6/21). Thematic analysis (Table 1) identified lack of emotional realism as the most frequently reported limitation. Despite this, 13 of 25 (52%) supported the integration of a formally structured AI history-taking practice into the curriculum, 8 (32%) were unsure, and 4 (16%) opposed.

Table 1. Perceived limitations of AI for history-taking practice. Of 21 users, 17 provided free-text limitations. Counts represent respondents per theme; respondents could contribute to multiple themes.

Theme	Respondents, n	Quote
Lack of emotional realism/human nuance	5	"It doesn't capture human emotions"
Non–curriculum-specific content	4	"[N]ot uni specific and usually less info than we need to know"
Inaccurate or superficial feedback	3	"Improper feedback"
Volunteers information	3	"It just gives you all the info"
Practical barriers	2	"Typing is more slow than speaking"
Repetitive scenarios	2	"[S]cenarios are too similar"

Discussion

Principal Findings

The key observation was the contrast between high initial uptake of AI for history-taking practice but limited sustained use, with students ranking it last below every other resource assessed. Reasons for this pattern remain uncertain. Purpose-built AI platforms have shown improved history-taking performance and confidence [4-6], possibly reflecting the benefit of specific tool design, although our study did not directly compare these approaches.

Ranking AI below passive resources such as static online notes indicates lower perceived usefulness. However, this may also reflect the social nature of peer role-play: reciprocal role-switching, observational learning, and immediate interpersonal feedback are inherent to peer practice but absent from solitary AI interaction. Even a tailored AI model is unlikely to replicate this social dimension, suggesting a complementary rather than substitutive role for peer learning.

Limited formal AI training may contribute to these perceptions. The chatbot used by Holderried and colleagues [4] required iterative prompt engineering, suggesting prompts instructing a model to withhold information or align with curriculum-specific scoring rubrics could plausibly address concerns about volunteered information and generic content. However, prompting strategies and training effects were not directly assessed in our study.

The most frequently cited limitation was a lack of emotional realism. Holderried and colleagues [4] also identified sounding human as one of two aspects requiring iterative prompt engineering, and despite otherwise positive usability ratings, around one-third of students found the chatbot too robotic.

Despite a 68% response rate and efforts to encourage participation irrespective of AI use, self-selection and nonresponse bias remain possible. The small, single-institution convenience sample renders these findings exploratory and limits generalizability. Findings reflect self-reported behavior and perceptions; objective performance was not assessed. However, limited formal AI training [7] and concerns about human nuance [4] align with previous research. Future multicenter studies should observe how students use these tools and whether this use affects objective OSCE performance.

Conclusions

Students prefer established resources and human interaction over generic AI use for history-taking practice, citing poor emotional realism and inaccurate feedback. Most had tried AI but few continued using it. Limited formal AI training is a possible but untested explanation, warranting further study. Interest in formally integrating a structured AI history-taking tool into the curriculum suggests students may see its value beyond their current self-directed use.

Generative AI Use

The authors declare that generative AI was not used in the preparation of this manuscript.

Funding

The authors declared no financial support was received for this work.

Data Availability

All data is available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Microsoft Forms questionnaire.
[PDF File (Adobe File), 353 KB-Multimedia Appendix 1]

Checklist 1

CHERRIES checklist.
[DOCX File (Microsoft Word File), 19 KB-Checklist 1]

References

1. Keifenheim KE, Teufel M, Ip J, et al. Teaching history taking to medical students: a systematic review. BMC Med Educ. Sep 28, 2015;15:159. [doi: 10.1186/s12909-015-0443-x] [Medline: 26415941]
2. Outcomes for graduates summary. General Medical Council. 2018. URL: https://www.gmc-uk.org/education/standards-guidance-and-curricula/standards-and-outcomes/outcomes-for-graduates/outcomes-for-graduates
3. Williams KN, Ramani S, Fraser B, Orlander JD. Improving bedside teaching: findings from a focus group study of learners. Acad Med. Mar 2008;83(3):257-264. [doi: 10.1097/ACM.0b013e3181637f3e] [Medline: 18316871]
4. Holderried F, Stegemann-Philipps C, Herschbach L, et al. A generative pretrained transformer (GPT)–powered chatbot as a simulated patient to practice history taking: prospective, mixed methods study. JMIR Med Educ. Jan 16, 2024;10:e53961. [doi: 10.2196/53961] [Medline: 38227363]

5. 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. Aug 16, 2024;10:e59213. [doi: 10.2196/59213] [Medline: 39150749]
6. Luo MJ, Bi S, Pang J, et al. A large language model digital patient system enhances ophthalmology history taking skills. NPJ Digit Med. Aug 4, 2025;8(1):502. [doi: 10.1038/s41746-025-01841-6] [Medline: 40760042]
7. Busch F, Hoffmann L, Truhn D, et al. Global cross-sectional student survey on AI in medical, dental, and veterinary education and practice at 192 faculties. BMC Med Educ. Sep 28, 2024;24(1):1066. [doi: 10.1186/s12909-024-06035-4] [Medline: 39342231]
8. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. Sep 29, 2004;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
9. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol. Jan 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
LLM: large language model
OSCE: objective structured clinical examination

Edited by Javad Sarvestan; peer-reviewed by Andrej Thurzo, Sara Kivity, Zhengdong Wu; submitted 31.Jul.2026; initial comments to author 02.Oct.2026; final revised version received 28.Oct.2026; accepted 11.Nov.2026; published 27.Nov.2026

Please cite as:
Goyal K, Noori S, Moyo S, Malhan NK
Self-Directed Use of General-Purpose AI Tools for History-Taking Practice Among Clinical-Phase Medical Students: Cross-Sectional Survey
JMIR Form Res 2026;10:e108381
URL: https://formative.jmir.org/2026/1/e108381
doi: 10.2196/108381

© Krishan Goyal, Shahmeer Noori, Sandiso Moyo, Neeraj Kumar Malhan. Originally published in JMIR Formative Research (https://formative.jmir.org), 27.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on https://formative.jmir.org, as well as this copyright and license information must be included.