ARTIFICIAL INTELLIGENCE-BASED TEACHING OF MEDICAL ENGLISH: OPPORTUNITIES AND LIMITATIONS

Djalilov Ikboljon Ikramovich
Lecturer, Department of Language Teaching in Medecine
Andijan Branch of Kokand University
E-mail: djalilovikboljon1968@gmail.com

Abstract: *This article examines how artificial intelligence can support the teaching of Medical English within an ESP framework and identifies the pedagogical risks that accompany its rapid adoption. Using a focused review of recent studies in language education and medical education, the paper analyzes opportunities for personalization, vocabulary development, simulated clinical communication, and assessment support, while also discussing hallucination, ethical concerns, overreliance, teacher training gaps, and infrastructural constraints in higher medical education.*

Keywords: *artificial intelligence; medical English; English for Specific Purposes; medical education; generative AI; ChatGPT; vocabulary learning; clinical communication; personalized learning; assessment; ethics.*

Introduction

Medical English is a branch of English for Specific Purposes (ESP) in which instruction is organized around professional discourse, authentic texts, and workplace communication. Medical students need more than general fluency: they must master terminology, read research, write case-based documents, and communicate accurately with patients and colleagues. For this reason, needs analysis is central to course design, because ESP should reflect real academic and professional demands [Hidayati & Meisani, 2023, pp. 467-468]. The rapid spread of artificial intelligence (AI), especially generative tools, has introduced new possibilities for teaching this kind of language.

Recent literature shows that AI can support speaking, writing, reading, pedagogy, and self-regulated learning, although it also brings technological and linguistic challenges [Crompton et al., 2024, p. 2503]. In medical education, AI systems have already been used for question answering, language editing, simulated patients, history-taking practice, and case-based materials [Wong, 2024, pp. 178-179]. These affordances are directly relevant to Medical English because they allow students to practice terminology and clinical communication repeatedly and at low cost.

Yet the medical context also amplifies the risks. AI may generate false information, fabricated references, or oversimplified responses; it may also encourage overreliance on machine output [Wong, 2024, pp. 179-180]. Reviews in healthcare education further emphasize concerns about plagiarism, privacy, transparency, legal accountability, and the

erosion of critical thinking [Sallam, 2023, pp. 1, 4]. Therefore, AI should be treated neither as a threat to be rejected nor as an authority to be trusted automatically. It must be evaluated as a pedagogical instrument whose value depends on human oversight and ethical regulation.

This article examines the opportunities and limitations of AI-based teaching of Medical English. It argues that AI can improve personalization, vocabulary learning, and simulated communication when used within an ESP framework, but its benefits remain conditional on task design, teacher mediation, and source verification.

Methods

This paper employs a focused qualitative literature review. Seven peer-reviewed publications from 2023 to 2025 were selected because they addressed AI in language education, AI in medical education, or Medical English/ESP curriculum design. The corpus included one study on medical-school ESP needs analysis, three reviews on AI in language or education, two reviews on AI in medical education, and one empirical study on AI-supported medical vocabulary learning in nursing education.

The sources were analyzed through thematic coding. Five categories guided the review: alignment with ESP needs, personalization and learner autonomy, vocabulary and clinical communication, teacher support and assessment, and pedagogical, ethical, and infrastructural limitations. The aim was not meta-analysis, but interpretive synthesis. This approach is suitable because research on AI specifically for Medical English is still emerging and often appears within broader studies of language teaching or medical education [Law, 2024, p. 1].

Results

AI and Alignment with ESP Needs

The literature indicates that AI is most useful when linked to authentic ESP outcomes rather than to generic conversation practice. Hidayati and Meisani show that medical ESP requires authentic materials, discipline-specific rhetorical features, and courses based on stakeholder needs [Hidayati & Meisani, 2023, pp. 467-468]. AI can respond to this demand by generating patient histories, referral letters, discharge instructions, case summaries, and article abstracts. Kasneci et al. note that large language models can work with domain-specific language and support texts relevant to medicine, including technical reports and records [Kasneci et al., 2023, pp. 2-3]. Law similarly reports that generative AI can provide immediate feedback on grammar, vocabulary, sentence structure, and rephrasing [Law, 2024, pp. 4-5]. These functions help align classroom tasks with real medical genres.

Personalization, Vocabulary Learning, and Clinical Communication

A major strength of AI is personalization. Law's review connects generative AI with rapid responses, individualized explanations, and greater learner autonomy [Law, 2024, pp. 4-5]. Crompton et al. also associate AI-supported language learning with self-regulation, goal setting, and reduced anxiety in speaking activities [Crompton et al., 2024, pp. 2521-2522]. In Medical English, where many students hesitate when using technical vocabulary orally, such support can lower barriers to practice.

Wong reports that AI may assist non-native English users through language editing and simulated patient dialogue or history-taking [Wong, 2024, pp. 178-179]. This is important because clinical communication involves not only lexical accuracy but also clarity, politeness, and empathy. AI chatbots can offer repeated low-stakes rehearsal of doctor-patient exchanges, enabling students to ask follow-up questions, clarify symptoms, and reformulate information in accessible language.

Direct evidence for vocabulary learning appears in Aboulfotoh's quasi-experimental study of nursing students using the Poe AI chatbot for English medical vocabulary. The experimental group improved from a mean pre-test score of 8.30 to a post-test score of 29.47, with a very large effect size (eta squared = 0.92) [Aboulfotoh, 2025, p. 155]. Learner interest was also much higher in the experimental group than in the control group [Aboulfotoh, 2025, p. 156]. Although this study focused on nursing students, it suggests that AI can strengthen terminology acquisition and motivation in medically oriented English classes.

Teacher Support and Assessment

The reviewed studies also show that AI can assist teachers. Kasneci et al. describe its value for individualized feedback, content creation, and instructional differentiation [Kasneci et al., 2023, p. 3]. Law adds that generative AI may reduce workload through support for lesson plans, quizzes, and some grading preparation [Law, 2024, p. 5]. Wong likewise notes that educators can generate case scenarios and formative questions with AI [Wong, 2024, p. 179]. However, evidence on assessment remains less developed. Crompton et al. argue that more innovation is still needed in the assessment dimension of AI-supported language education [Crompton et al., 2024, p. 2520]. At present, AI appears more credible for formative support than for high-stakes evaluation.

Limitations and Risks

The same literature identifies serious limitations. First, reliability remains unstable. AI may provide inaccurate, misleading, or fabricated information [Wong, 2024, pp. 179-180]. Sallam also identifies hallucination, misinformation, and incorrect citations as recurrent concerns in healthcare education [Sallam, 2023, pp. 1, 4]. In Medical English, such errors can lead students to internalize unsafe or inaccurate language.

Second, overreliance can weaken higher-order learning. Law warns that constant dependence on AI feedback may reduce critical thinking [Law, 2024, pp. 4-5]. Wong argues that AI should not replace clinical reasoning [Wong, 2024, p. 181], and Kasneci et al. insist on human oversight so that learners continue to develop judgment and evaluative skills [Kasneci et al., 2023, p. 6].

Third, ethical and institutional problems remain unresolved. Privacy, plagiarism, authorship, transparency, and bias are highlighted across the literature [Kasneci et al., 2023, pp. 7-8; Sallam, 2023, p. 4]. These issues are especially sensitive in medical contexts, where prompts may contain patient-related information. Effective implementation therefore requires policy support and teacher training. Kasneci et al. call for educator guidance, while Law

stresses stakeholder engagement and continuing professional development [Kasneci et al., 2023, p. 7; Law, 2024, p. 12].

Discussion

The findings suggest that AI is most effective in Medical English when used as a scaffold for ESP performance, not as a substitute for expert instruction. Medical English is not simply general English plus technical vocabulary; it requires genre awareness, precision, and pragmatic sensitivity. For that reason, the success of AI depends less on novelty than on instructional design.

Several practical applications follow from the reviewed literature. Before class, AI can generate glossaries, flashcards, and simplified readings. During class, it can support role-play prompts, simulated patient interviews, case presentations, referral letters, and collaborative editing of clinical documents. After class, it can provide paraphrasing practice, comprehension questions, and draft feedback. These uses remain compatible with the ESP principle that language tasks should reflect authentic professional communication [Hidayati & Meisani, 2023, pp. 467-468].

At the same time, AI use must be paired with structured validation. Students should compare chatbot output with textbooks, teacher feedback, clinical guidelines, and peer discussion. In this way, AI-generated text becomes material for analysis rather than a final answer. A useful pedagogical cycle is AI-supported drafting, evidence-based verification, and reflective revision. This model preserves efficiency while protecting academic standards.

Assessment should also be redesigned. If students can outsource parts of writing to AI, then teachers should evaluate process as well as product. Prompt logs, source checks, revision commentaries, and short oral defenses can make AI use transparent and accountable. Such measures help students develop the critical literacy needed to detect hallucinations, bias, and formulaic language.

Finally, institutional policy is essential. Medical universities need clear rules on acceptable AI use, confidentiality, citation disclosure, and the boundary between assistance and misconduct. Staff development is equally important, since teachers who do not understand AI cannot guide students effectively [Kasneci et al., 2023, pp. 7-8; Law, 2024, p. 12].

Conclusion

AI offers significant opportunities for teaching Medical English within an ESP framework. It can personalize learning, strengthen vocabulary acquisition, support simulated clinical communication, and assist teachers with materials and formative assessment. However, these benefits are limited by unreliable outputs, overreliance, privacy concerns, plagiarism risks, and uneven institutional readiness. AI should therefore be integrated as a supervised educational assistant rather than as an autonomous authority. Future research should examine its long-term effects on communication accuracy, disciplinary reasoning, and safe classroom practice in medically oriented language education.

References

1. Aboulfotoh, M. A. (2025). Integration of Poe AI chatbot into English medical vocabulary learning classes for students of nursing. International Journal of Research in Educational Sciences, 8(4), 103-171. https://doi.org/10.29009/ijres.8.4.3

2. Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. British Journal of Educational Technology, 55, 2503-2529. https://doi.org/10.1111/bjet.13460

3. Hidayati, L., & Meisani, D. R. (2023). The needs analysis of English for specific purposes: A study in an Indonesian medical school. Indonesian Journal of English Language Teaching and Applied Linguistics, 7(2), 465-479.

4. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

5. Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. Computers and Education Open, 6, 100174. https://doi.org/10.1016/j.caeo.2024.100174

6. Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. Healthcare, 11(6), Article 887. https://doi.org/10.3390/healthcare11060887

7. Wong, R. S.-Y. (2024). ChatGPT in medical education: Promoting learning or killing critical thinking? Education in Medicine Journal, 16(2), 177-183. https://doi.org/10.21315/eimj2024.16.2.13