**Author**

Dr. Odebode, Olusegun

**Project title**

Artificial Intelligence, when used responsibly, can serve as a valuable tool in enhancing core competencies in the Family Medicine Residency Program

**Supervisor**

Dr Conyers, Roop

**Site**

Annapolis Valley

**Project type**

Position Paper

This project followed the guidelines for 2024/2025 for Dalhousie Family Medicine

**Date**

February 6th, 2026

# ARTIFICIAL INTELLIGENCE, WHEN USED RESPONSIBLY, CAN SERVE AS A VALUABLE TOOL IN ENHANCING CORE COMPETENCIES IN THE FAMILY MEDICINE RESIDENCY PROGRAM.

## Introduction

Machines that use human intelligence to solve problems, from simple to complex, have existed for a long time. The term "Artificial Intelligence (AI)" was first coined by John McCarthy during the Dartmouth conference in 1956 (1). Since then, research has focused on various components of AI, including rule-based systems, machine learning (ML), deep learning (DL) algorithms, and, more recently, natural language processing (NLP) (2,3). This evolution has enabled AI to address issues that directly or indirectly impact our daily lives (4). For instance, AI was used to break wartime codes in 1942, Deep Blue (a supercomputer) defeated chess champions, and technologies like Google Maps assist us with navigation. More recently, AI has been instrumental in taking medical notes and even diagnosing medical conditions (eg, Heidi, ChatGPT, OpenEvidence, FullCode). Beyond this, AI could help improve knowledge, especially in clinical reasoning, among medical learners and residents in family medicine practice (5).

Institutions have begun incorporating the teaching of Generative AI into their postgraduate family medicine education (6). Similarly, in the context of continuous training in family medicine residency, universities such as Dalhousie's have adopted competency-based training that emphasizes clinical reasoning, selectivity, communication, professionalism, and procedural skills. This position paper will emphasize clinical reasoning, as it is essential for translating theoretical medical knowledge into everyday patient care (7,8)

The most widely used tools for assessing competency in Canadian FM residency programs are the Field Note and preceptor feedback. Recent data indicate that residents are increasingly focusing on clinical reasoning when completing their field notes (9). Often, self-reported field notes or preceptors' feedback can

be vague; as such, virtual AI models to support training and provide real-time feedback are argued for and against the conventional approach (4).

Currently, it is not common practice to use AI tools to support residents in tracking their clinical reasoning in areas such as generating hypotheses and differential diagnoses, gathering focused data (history and physical examination), interpreting data, making clinical decisions, and setting goals and objectives across a patient's life cycle. Some fields of medicine have demonstrated the benefits of AI use, including AI-based emotional recognition to improve students' academic performance(10) and AI-based simulation to improve communication among pediatric residents (11). Further, other specialties, such as radiology, orthopedics, and internal medicine, have incorporated AI into residency training(12,13).

**Background and Rationale**

Clinical reasoning development is essential for real-world medical care. Clinicians employ different models of cognitive reasoning when interacting with patients in various situations. Examples of these models include heuristics (fast and frugal), pattern recognition (illness scripts), clinical algorithms (protocol-based), and the hypothetico-deductive approach. Among these, AI has been suggested as a tool to help residents improve their use of the clinical algorithm (7) and the illness script model (14).

The amount of medical information available to residents is expected to double every 73 days (15). At the time of writing this position statement, Dalhousie University had incorporated Full Code, an AI-driven medical simulation application, into family medicine residency training. However, it has since been removed from the program, primarily due to sponsorship issues rather than a lack of popularity among residents. Similar applications have been cited in other journals (5) and can be used in diverse ways to improve residents' clinical reasoning. When used strategically, such AI-powered simulations can not only improve learners'

performance given the vast knowledge they are expected to have but also enhance patient care and outcomes (16).

This paper aims to provide a blueprint for effectively using such tools to strengthen residents' core competency in clinical reasoning, especially in areas where they seek increased exposure. Limited exposure may result from the rarity of some instances, teaching sites that seldom offer those learning opportunities, or the varied medical exposure of internationally trained physicians compared to Canadian-trained physicians. In addition, residents with busy schedules may prioritize acute and emergent patients, leaving little time to reflect on cases that could be improved (5).

These areas of need range from common but diagnostically challenging presentations to rare, high-stakes cases, and include situations where learners require tailored learning plans or extended training programs. By leveraging AI, programs can help prevent conditions that lead to prolonged training with limited opportunities to master specific cases or procedures.

Various fields are working to make feedback more objective and to foster a safe environment for learners to conduct self-assessments (8). Embedding AI in Electronic Medical Records (EMR) to assess clinical reasoning in real time can create a secure environment that enables learners to identify and rectify deficiencies quickly. This approach reduces the subjectivity of preceptors' assessments—which, although not inherently flawed, can often be perceived by learners as delayed and not reflective of their competence—and enhances the training experience.

First-year Residents at this stage of their training are striving to build confidence while managing anxiety about feedback from preceptors, which could significantly impact their learning experience(9) (17). The emotional component of learning is well documented in the work of Vistorte AOR (10), and evidence indicates

that AI can provide personalized, emotionally supportive feedback. Due to feedback-related anxiety, residents may unintentionally avoid documenting feedback on complex or challenging cases, even though these cases are essential for developing robust clinical reasoning skills (9). This gap presents an opportunity for AI-driven integration of field notes within the EMR, enabling the systematic capture of observed feedback in a learner-centred manner.

AI-based tools can evaluate clinical reasoning at both the individual and collective levels, offering an assessment scale that exceeds human capability. This technology enables a deeper understanding of and improvements to the Clinical Learning Environment and its impact on residents' documentation of clinical reasoning practices, ultimately strengthening the training process in family medicine (18).

In alignment with the guiding principles established by institutions such as Dalhousie University, AI is intended to enhance rather than replace critical thinking (19). This position paper expands on ways AI can be effectively utilized in this role. To support this position, extensive research was conducted using databases such as PubMed, Google Scholar, and Web of Science. This research included university policy statements on Generative AI (19) and position papers from other specialties(20). Document selection criteria included studies that explicitly mentioned Artificial Intelligence (e.g., AI, medical education, and core competencies) in headings or MeSH terms. Articles published before 2015 were excluded, given that earlier technologies are now largely considered obsolete.

**Evidence Supporting AI in Medical Education**

1. **Comparative Performance:** AI has been shown to possess superior medical knowledge and to demonstrate clinical reasoning. It can assist and enhance clinical reasoning among Residents. In a cited rheumatology study, ChatGPT -4 (Chat Generative Pre-trained Transformer-4) demonstrated superior clinical reasoning scores on the R-IDEA scale compared with junior residents at various

stages of diagnosis and management (first- and second-line management) (17). Another comparative study compared this Large language model (LLM), ChatGPT, with an older version on the American Board of Family Medicine's In-Training Examination (ABFM-ITE). The comparison showed that the newer AI version can improve performance. The performance's scalability was remarkable, increasing from 56% to 86%, a substantial 410-point improvement. This was displayed by ChatGPT 4 (17) (21). We are currently using ChatGPT-5. Notably, it was able to refine its wrong answer without being updated with the correct answers.

The aforementioned capability can be used as a teaching tool and to maximize learning. Experts and attending staff are trained throughout their careers to teach and provide feedback to learners. Significant time and resources are also necessary to ensure residents are fully aware that the attending still has other responsibilities, such as caring for their own patients, in addition to their role. This gap can be filled or at least supported by AI. For example, AI is trained using a set of databases. The advancement of generative AI has taken it further; this innovation can access multiple reliable internet resources and additional databases to enhance its knowledge and provide information to residents.

2. **Attitudes and Acceptance of AI among Clinicians and Residents:** Acceptance of AI tools among preceptors and family physicians has increased in the past decade (22). A 2019 international Delphi consensus study found that healthcare providers in family medicine viewed AI with both caution and optimism, recognizing its potential to support decision-making while expressing concerns about transparency, bias, workflow integration, and overreliance. The study highlighted the importance of strict governance and clinician oversight, offering moderate-quality expert consensus evidence rather than direct data on effectiveness (23).

   A recent mixed-methods study was generally positive about AI use. However, it showed that AI did not significantly enhance clinical decision-making and, in some cases, had a negative impact.

Nonetheless, physicians reported that simulation-based decision support improved clinical decisions and believed AI could further aid decision-making through more informed human-computer interfaces(24). This study also discussed proposals for physician training in AI use.

A subsequent qualitative descriptive study, utilizing semi-structured interviews and thematic analysis, found that AI helps physicians free up time, enabling more meaningful and relevant patient discussions while conversations are transcribed (23). Notably, 60% of residents expressed interest in using AI as a supportive tool, with acceptance contingent upon trust in AI's accuracy and the EMR output. Recent work by Daryoush Shahin emphasizes that residents' understanding and application of AI increase confidence, leading to more efficient, higher-quality care (15).

Given the expansive scope of Family Medicine—which encompasses chronic disease management, acute care, and psychosocial issues—AI can provide structured reasoning support across a diverse range of clinical scenarios (24,25). Furthermore, AI has the potential to enhance clinical judgment in the early years of residency, when uncertainty tends to peak (15(25)). For example, residents may struggle to develop treatment plans (particularly when making decisions in the presence of multiple comorbidities), may become overwhelmed in high-pressure situations due to limited experience, or may find it challenging to apply clinical guidelines to unique patients (17).

**Core Competencies Enhanced by AI**

1. ***Clinical Reasoning:*** AI supports hypothesis generation, differential diagnosis, data interpretation, and management planning. It assists in decision-making and helps set goals and objectives, complemented by expert feedback (26). According to the CFPC evaluation objectives, traditional methods to enhance this skill set include field notes, hypothesis-generation topics, illness scripts, and chart reviews. When using AI—whether via the EMR, simulations, or open-source tools—it provides a real-time list of differentials, related questions to assist with history-taking, illness scripts,

potential physical examinations, management strategies, investigations, imaging, treatment options, referrals, and follow-up timelines for each differential. The list is ordered by relevance. Some applications even display their reasoning process for the listed differentials, which can help learners adjust their decisions to the patient's context. This is similar to most case preparations, except that with the use or implementation of AI, information is easier to access and consolidate. Over time, repeated exposure to such structured reasoning can be internalized by the learner.

a. **AI-driven EMR and Field note:** Field notes have been an essential means for program directors to assess family medicine residents, using the In-Training Evaluation Reports (ITER) for periodic reviews. One proposed approach is an AI-integrated Electronic Health Record (EHR) capable of assessing documentation for clinical-reasoning field notes(8) (27). In Schaye's work, the accuracy of the documentation was not evaluated; however, the emphasis was on documentation and the quality of the documented clinical reasoning, a vital skill learners must develop(8). Another study demonstrated that AI on EHRs can assess the impact of the learning environment on clinical reasoning documentation(8,27), such as Internal Medicine residents documenting their clinical reasoning more effectively during the early hours of a night shift than during a busy daytime shift. This shows that learners can assess their timely progress in the EMR to address concerns their preceptor identified and see how the schedule or time of day affects their documentation. When residents review and reflect on the AI- generated comments,

   they can validate and adopt necessary changes for improvement. AI feedback can be programmed to run daily, and learners can see their progress over time (25), in time for their following periodic review.

b. **Using AI to Identify Residents' Success in the CCFP Exams:** In Abhisht's work at the University of Ottawa, AI was used retrospectively to identify learners with a high failure rate (25). By analyzing ITER data and incorporating preceptor comments, a multimodal approach employing quantitative,

qualitative, and explainable AI achieved a predictive success rate of 89%, which exceeds that of previous methods without explainable AI (25,28) . This approach aims to identify areas of weakness early in PGY-1 training, enabling the development of targeted learning plans.

The study hypothesized that residents who required discussion of program objectives had a higher failure rate, whereas those identified as team players had a higher success rate. Learning plans can be implemented after the first or second review, and their implementation should be viewed positively rather than as a drawback, as ensuring that all learners have one may be beneficial (26).

c. **AI simulation tools:** The use of simulations and virtual reality to enhance core competencies is widely adopted in most medical institutions. However, AI-driven simulation with feedback is still in its early stages. A German trial demonstrated improvements in clinical decision-making, particularly in creating context and gathering relevant information about the patient's case, among medical learners exposed to an AI-simulation with feedback compared with those who experienced only AI-simulations without AI-generated feedback (29). This approach fosters a conducive learning environment with real-time feedback. Learners can practise privately, repeatedly, to achieve mastery. As noted earlier, conventional methods or expert feedback can be vague, leading learners to overlook correct actions or connections. AI-driven feedback can highlight these omissions and be specific to a particular competency. A recent study involving clinically naive medical learners showed that their clinical reasoning skills improved with AI-generated feedback on clinical reasoning assessments (11).

2. **Other Core Competencies with AI input:** AI tools can model patient-friendly explanations, thus supporting residents' communication training (11,30). In most fields, the use of simulations to improve communication is more evident; the order in which questions should be asked can be practised at the learner's convenience and privacy. This eliminates the awkwardness of correcting minor errors learners might perceive as negligible, even though they are vital to the patient's

encounter. As with good communication, the sequence of retrieving information from a patient can boost or reinforce trust and establish a sense of being well cared for.

Other core competencies, such as selectivity, which is relatively new and challenging to assess (31), can be addressed using AI to prioritize investigations and treatments by evaluating evidence-based options (32). In the area of professionalism, AI reduces cognitive load and burnout, allowing residents to focus on reflective practice. These competencies can be assessed using AI-generated feedback, given the demonstrated superiority of AI clinical knowledge.

**Counterarguments and Considerations**

Like many practical tools, AI can significantly assist trainees, preceptors, and family medicine faculty in enhancing the competencies discussed. However, the extent of its use will determine any potential harm it may cause (14,33).

1. **Overreliance Risk:** There's a risk that residents might depend too heavily on AI, weakening their independent reasoning in other resource-depleted environments(12). This can be mitigated by the Generative AI guidelines and policies, which require all residents to review AI-generated documentation before signing off, ensuring they remain the sole authors and are therefore responsible for the content (19). To explain, as residents review AI-supported documentation, they are exposed to how diagnoses are constructed, which red flags are considered, and what management strategies are proposed. Over time, this supports both passive and active learning, enhancing knowledge and written communication in a relatively safe environment. AI can constantly provide prompts for improvement. For example, a physical examination that is missing will be recorded as "physical not done." The next time, learners might be more inclined to remember to do the necessary examination. As mentioned earlier, the clinical reasoning feedback can be reviewed by the learner. Conventionally, residents have been observed to document their successes more

frequently and tend to avoid noting down areas for improvement. This may create a bias toward focusing on their strengths. The reasons for this behaviour could stem from a desire to avoid appearing deficient or a tendency to reinforce positive experiences (9). Regardless, this highlights the potential for AI to assist in documentation, possibly through an EMR accessible only to students. When a preceptor points out areas for improvement during discussions or chart reviews, it could serve as confirmation rather than a surprise.

2. **Errors and Inaccuracies:** In family medicine, the use of open-source tools, such as ChatGPT, has demonstrated high performance on in-training exams, showcasing the potential for learners to enhance their decision-making abilities (21). However, recent studies in a British journal raised concerns about the limitations of ChatGPT 4.0, particularly in emergency department contexts for ECG interpretation, where it exhibited a 46% error rate (34). More research is needed to clarify when and how AI can safely assist even experienced clinicians. As human experts in the field of medicine, there are daily learning opportunities given the rate at which information doubles every few months(15). This means that both humans and AI will continue to have knowledge gaps (2). The notion is not an excellent tool, but can it be objective and reliable enough to improve learners' reasoning in everyday practice (5).

   To address these limitations further, new AI tools are designed to factor in medical cases, with access to evidence-based resources. For example, the recently released AI tool Deepseek was tested against a group of critical care residents using PubMed, Medscape, and UpToDate. The AI-assisted group excelled in completeness, clarity, and usefulness, leading to a significant improvement in residents' diagnostic performance and a significant enhancement in differential diagnosis (35).

   Other barriers to the use of AI have been identified, including inaccuracies (36). Newer models of AI chatbots are designed to reduce inaccuracies or hallucinations (a phenomenon in which the information generated is statistically possible but inaccurate) (16,37). However, there are medically

designed AIs that can be relied upon and held to the same standards as widely accepted resources, such as OpenEvidence, which, when asked a question, provides answers with cited sources and evidence strength when available.

3. **Bias:** AI is only as reliable as its training database, which can introduce bias (38). However, these databases are from current and existing literature. Hence, its bias will not be considered much higher than that of conventional textbooks or other commonly accepted medical websites such as UpToDate. Moreover, with medically endorsed AI, reference prompts can be used to guide the reliability of the information source, such as in clinical applications that must prioritize patient confidentiality and responsible integration (19,26). Resident training on AI, incorporating AI curricula into family medicine residency training, can significantly address these concerns. For future consideration, AI-trained medical learners will recognize this bias and make decisions that address it(6). Other problems, such as ethical and privacy issues, can also be addressed by educating students about AI(4).
4. **Insufficient training, Digital literacy, and Resistance to Change:** Another barrier is that most clinicians do not have sufficient training or understanding of how AI operate. The risk is that learners might not critique AI-generated content effectively enough to realize its potential to improve clinical reasoning. Also, there might be an inconsistency in how it is used to advance clinical reasoning. For AI to be considered a valuable tool for improving decision-making, learners and preceptors must understand the basic principles of AI (14). This knowledge remains limited (14,39). A practical approach would be to teach both residents and staff the responsible use of AI. Residents must be taught critical appraisal, real-time performance monitoring, and digital literacy (6,26).

   In addition to integrating an AI-driven clinical reasoning tool, it is vital to have a framework that educates residents about AI. This can be achieved through incorporating AI-assisted learning modules within family medicine residency curricula, using the AIFM curriculum model as a guide (6),

Conduct specialty-specific studies to assess AI's effect on FM residents (40), expand on the ethical guidelines for responsible use in training and patient care (20,26), train faculty and residents on the effective use of AI in clinical reasoning (16,26).

Resistance to change presents another challenge. Surprisingly, more physicians, when exposed to simulation-based applications such as Full Code scenarios adopted at the program level, can be convinced of their usefulness. These simulations serve as valuable resources to bridge gaps in learning and areas of interest, especially when exposure is limited (5). For example, simulations covering the treatment of Addison's crisis or frostbite can be conducted at learners' convenience, allowing them to discuss these cases at the end of a day clinic. With repeated practice, learners may gain confidence and feel less anxious in real-life situations. PBL addresses this, but with an AI-driven approach, all residents actively participate.

**Policy makers' consideration**

The integration of AI to enhance clinical reasoning should align with the responsible use guidelines outlined by Dalhousie University and the CFPC in their statement on AI for family medicine (19,20). The use of AI scribes by learners ought to be at the discretion of their preceptors (19). However, situations may arise where a preceptor is hesitant to adopt AI; even without an explicit prohibition, a lack of encouragement may be perceived as rejection, potentially discouraging learners' initiative. To address this, residents and preceptors should agree on assessment methods. Preceptors are encouraged to initiate early discussions about AI use, or learners should be empowered to make responsible decisions regarding AI without necessarily seeking preceptor approval. Incorporating an AI curriculum is vital, as it establishes AI usage as a recognized requirement for both learners and staff.

Considering the importance of responsible AI use, it is valuable to include an AI curriculum in residency training, as suggested by the University of McGill (6) and the Dalhousie AI policy statement(19). Learners

should be instructed on accountable AI use through a curriculum that begins in medical school and continues during residency. This approach can further help develop their clinical reasoning skills, allowing confident users to leverage AI as a supportive resource (6).

Finally, fostering collaboration with technologists to design and refine AI solutions can improve performance and acceptance in family medicine (38). Although this process may require significant investment of time and resources, involving approvals and committee reviews, a thoughtful approach guided by research could yield valuable outcomes. It is crucial to determine whether we have an AI solution that supports medical learners, ensuring decisions are made carefully and thoughtfully.

**Recommendations**

1. AI should be considered as a teaching and learning tool for residents and medical learners (16,17,29).
2. AI curriculum should be incorporated into family medicine residency programs (6,17,20)
3. AI should be integrated into the EMR for resident monitoring and assessment in addition to its current use as a documentation tool (8,19,27).
4. AI should be used as a tool for field note-taking. The preceptor remains the final assessor of trainees in certified competencies. The preceptor can use an AI cumulative assessment to track the trainee's progress within a learning plan(27)
5. AI should be used to identify learners who require additional support in areas of limited exposure and ultimately succeed in the CFPC exams (25,28).
6. The family medicine program should collaborate with the technologists create an EMR with AI-driven feedback (20,27).

Family Medicine Residency programs should stay alert to advances in generative AI, which will significantly transform healthcare, and remain open to integrating these tools into patient care. Clinicians should critically assess the benefits of AI to improve their practice and adopt tools that genuinely enhance outcomes (19). However, it is essential to be aware that errors and omissions are common challenges with generative AI; understanding a new tool's performance and safety measures is vital (6).

## Conclusion

AI can be a valuable tool for enhancing core competencies in family medicine residency programs. Research has shown that AI can assess residents' documentation of clinical reasoning and analyze its relationship with the clinical learning environment(8). This is particularly relevant to clinical documentation and field note reflection. When used correctly and responsibly, AI-driven simulation applications should be used for practice and to reinforce learning (5).

Family medicine residency programs can use AI to improve resident training and support preceptors in providing valuable feedback to learners and other members of the education team (27). Maximizing the benefits of AI while maintaining professional standards can be achieved by integrating an AI curriculum into the residency program. It is essential to ensure transparency, confidentiality, and clinical oversight. Family medicine residency programs should consider using AI tools to enhance core competencies beyond improving documentation efficiency. This proactive approach can help prevent students from needing remediation and improve learners' competencies in core clinical reasoning skills (17).

**Declaration:** AI, Grammarly was used in editing this paper.

## References

1. Liaw W, Kakadiaris IA. Artificial Intelligence and Family Medicine: Better Together. Fam Med. 2020 Jan 3;52(1):8–10.

2. Matos T, Santos W, Zdravevski E, Coelho PJ, Pires IM, Madeira F. A systematic review of artificial intelligence applications in education: Emerging trends and challenges. Decis Anal J. 2025 Jun;15:100571.

3. Al-Zaiti SS, Alghwiri AA, Hu X, Clermont G, Peace A, Macfarlane P, et al. A clinician's guide to understanding and critically appraising machine learning studies: a checklist for Ruling Out Bias Using Standard Tools in Machine Learning (ROBUST-ML). Eur Heart J - Digit Health. 2022 Jul 6;3(2):125–40.

4. Alowais SA, Alghamdi SS, Alsuhebany N, Alqahtani T, Alshaya AI, Almohareb SN, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ. 2023 Sep 22;23(1):689.

5. Abdulnour REE, Parsons AS, Muller D, Drazen J, Rubin EJ, Rencic J. Deliberate Practice at the Virtual Bedside to Improve Clinical Reasoning. N Engl J Med. 2022 May 19;386(20):1946–7.

6. Tolentino R, Hersson-Edery F, Yaffe M, Abbasgholizadeh-Rahimi S. AIFM-ed Curriculum Framework for Postgraduate Family Medicine Education on Artificial Intelligence: Mixed Methods Study. JMIR Med Educ. 2025 Apr 25;11:e66828.

7. Ng IKS, Goh WGW, Teo DB, Chong KM, Tan LF, Teoh CM. Clinical reasoning in real-world practice: a primer for medical trainees and practitioners. Postgrad Med J. 2025 Jan 1;101(1191):68–75.

8. Schaye V, DiTullio DJ, Sartori DJ, Hauck K, Haller M, Reinstein I, et al. Artificial intelligence based assessment of clinical reasoning documentation: an observational study of the impact of the clinical learning environment on resident documentation quality. BMC Med Educ. 2025 Apr 22;25:591.

9. Zaki N, Cavett T, Halas G. Field note use in family medicine residency training: learning needs revealed or avoided? BMC Med Educ. 2021 Dec;21(1):451.

10. Vistorte AOR, Deroncele-Acosta A, Ayala JLM, Barrasa A, López-Granero C, Martí-González M. Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review. Front Psychol. 2024;15:1387089.

11. Merritt C, Glisson M, Dewan M, Klein M, Zackoff M. Implementation and Evaluation of an Artificial Intelligence Driven Simulation to Improve Resident Communication With Primary Care Providers. Acad Pediatr. 2022 Apr;22(3):503–5.

12. Krishna Nk R, R.S. A, K S. Artificial Intelligence in Radiology: Augmentation, Not Replacement. Cureus [Internet]. 2025 Jun 17 [cited 2025 Oct 31]; Available from: https://www.cureus.com/articles/378642-artificial-intelligence-in-radiology-augmentation-not-replacement

13. Goh E, Gallo R, Strong E, Weng Y, Kerman H, Freed J, et al. Large Language Model Influence on Management Reasoning: A Randomized Controlled Trial. medRxiv. 2024 Aug 7;2024.08.05.24311485.

14. Hallquist E, Gupta I, Montalbano M, Loukas M. Applications of Artificial Intelligence in Medical Education: A Systematic Review. Cureus. 2025 Mar;17(3):e79878.

15. Shahin D, Sostorecz S, Shariff A, Ng-Pellegrino A, Goodbred A. Role of Artificial Intelligence (AI) in Family Medicine. In: Artificial Intelligence in Medicine and Surgery - An Exploration of Current Trends, Potential Opportunities, and Evolving Threats, Volume 3 [Working Title] [Internet]. IntechOpen; 2025 [cited 2025 Dec 1]. Available from: https://www.intechopen.com/online-first/1220622

16. Alli SR, Hossain SQ, Das S, Upshur R. The Potential of Artificial Intelligence Tools for Reducing Uncertainty in Medicine and Directions for Medical Education. JMIR Med Educ. 2024 Nov 4;10(1):e51446.

17. Kayacan Erdoğan E, Babaoğlu H. Clinical Reasoning and Knowledge Assessment of Rheumatology Residents Compared to AI Models: A Pilot Study. J Clin Med. 2024 Dec 5;13(23):7405.

18. Çiçek FE, Ülker M, Özer M, Kıyak YS. ChatGPT versus expert feedback on clinical reasoning questions and their effect on learning: a randomized controlled trial. Postgrad Med J. 2025 Apr 22;101(1195):458–63.

19. ArtificialIntelligenceinResidencyTraining.pdf [Internet]. [cited 2025 Oct 7]. Available from: https://cdn.dal.ca/content/dam/dalhousie/pdf/faculty/medicine/departments/core-units/postgrad/ArtificialIntelligenceinResidencyTraining.pdf

20. Slade S. CFPC Statement on Artificial Intelligence.

21. Wang T, Mainous AG, Stelter K, O'Neill TR, Newton WP. Performance Evaluation of the Generative Pre-trained Transformer (GPT-4) on the Family Medicine In-Training Examination. J Am Board Fam Med. 2024 Jul 1;37(4):528–82.

22. Nash DM, Thorpe C, Brown JB, Kueper JK, Rayner J, Lizotte DJ, et al. Perceptions of Artificial Intelligence Use in Primary Care: A Qualitative Study with Providers and Staff of Ontario Community Health Centres. J Am Board Fam Med. 2023 Apr 3;36(2):221–8.

23. Liyanage H, Liaw ST, Jonnagaddala J, Schreiber R, Kuziemsky C, Terry AL, et al. Artificial Intelligence in Primary Health Care: Perceptions, Issues, and Challenges: Primary Health Care Informatics Working Group Contribution to the Yearbook of Medical Informatics 2019. Yearb Med Inform. 2019 Aug;28(01):041–6.

24. Hah H, Goldin DS. How Clinicians Perceive Artificial Intelligence-Assisted Technologies in Diagnostic Decision Making: Mixed Methods Approach. J Med Internet Res. 2021 Dec 16;23(12):e33540.

25. Joshi AM, Mortezaagha P, Inkpen D, Seale E, Archibald D, Noel K, et al. Early identification of Family Medicine residents at risk of failure using Natural Language Processing and Explainable Artificial Intelligence [Internet]. Medical Education; 2024 [cited 2025 Dec 5]. Available from: http://medrxiv.org/lookup/doi/10.1101/2024.12.07.24318566

26. Van Baalen S, Boon M, Verhoef P. From clinical decision support to clinical reasoning support systems. J Eval Clin Pract. 2021 Jun;27(3):520–8.

27. Schaye V, Guzman B, Burk-Rafel J, Marin M, Reinstein I, Kudlowitz D, et al. Development and Validation of a Machine Learning Model for Automated Assessment of Resident Clinical Reasoning Documentation. J Gen Intern Med. 2022 Jul;37(9):2230–8.

28. Ariaeinejad A, Patel R, Chan TM, Samavi R. P031: Using machine learning algorithms for predicting future performance of emergency medicine residents. Can J Emerg Med. 2017 May;19(S1):S88–S88.

29. Brügge E, Ricchizzi S, Arenbeck M, Keller MN, Schur L, Stummer W, et al. Large language models improve clinical decision making of medical students through patient simulation and structured feedback: a randomized controlled trial. BMC Med Educ. 2024 Nov 28;24(1):1391.

30. Potter L, Jefferies C. Enhancing communication and clinical reasoning in medical education: Building virtual patients with generative AI. Future Healthc J. 2024 Apr;11:100043.

31. Defining competence for the purposes of certification by the College of Family Physicians of Canada:

32. Knopp MI, Warm EJ, Weber D, Kelleher M, Kinnear B, Schumacher DJ, et al. AI-Enabled Medical Education: Threads of Change, Promising Futures, and Risky Realities Across Four Potential Future Worlds. JMIR Med Educ. 2023 Dec 25;9:e50373.

33. Wang M, Sun Z, Jia M, Wang Y, Wang H, Zhu X, et al. Intelligent virtual case learning system based on real medical records and natural language processing. BMC Med Inform Decis Mak. 2022 Dec;22(1):60.

34. Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. Drazen JM, Kohane IS, Leong TY, editors. N Engl J Med. 2023 Mar 30;388(13):1201–8.

35. Wu X, Huang Y, He Q. A large language model improves clinicians' diagnostic performance in complex critical illness cases. Crit Care. 2025 Jun 6;29(1):230.

36. Azamfirei R, Kudchadkar SR, Fackler J. Large language models and the perils of their hallucinations. Crit Care. 2023 Mar 21;27(1):120.

37. Han Z, Battaglia F, Udaiyar A, Fooks A, Terlecky SR. An explorative assessment of ChatGPT as an aid in medical education: Use it with caution. Med Teach. 2024 May 3;46(5):657–64.

38. Sutera P, Bhatia R, Lin T, Chang L, Brown A, Jagsi R. Generative AI in Medicine: Pioneering Progress or Perpetuating Historical Inaccuracies? Cross-Sectional Study Evaluating Implicit Bias. JMIR AI. 2025 Jun 24;4:e56891–e56891.

39. Cross J, Robinson R, Devaraju S, Vaughans A, Hood R, Kayalackakom T, et al. Transforming Medical Education: Assessing the Integration of ChatGPT Into Faculty Workflows at a Caribbean Medical School. Cureus [Internet]. 2023 Jul 5 [cited 2026 Jan 17]; Available from: https://www.cureus.com/articles/154384-transforming-medical-education-assessing-the-integration-of-chatgpt-into-faculty-workflows-at-a-caribbean-medical-school

40. Resident Physicians' Perceptions of Artificial ... | MedEdPublish [Internet]. [cited 2025 Dec 4]. Available from: https://mededpublish.org/articles/15-29