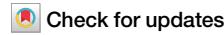




Implications of integrating large language models into clinical decision making



Michael Christof ¹ & Antonis A. Armoundas ^{2,3}

Clinical reasoning is essential for effective medical practice, requiring framing of the encounter, diagnosing, and managing patient care. Large Language Models (LLMs) have recently emerged as a tool to supplement and enhance clinical reasoning; in this Comment, we discuss their potential to support clinician-level clinical reasoning.

Clinical reasoning is the mental process through which clinicians gather, interpret, and integrate clinical information to create an abstract summary of a case, which is known as a problem representation¹. This process involves generating and refining diagnostic ideas, ultimately leading to a diagnosis. It utilizes both analytical reasoning and intuitive pattern recognition, with the latter grounded in experience-based “illness scripts” that link risk factors, underlying mechanisms, and clinical outcomes. Effective clinical reasoning is crucial for accurate diagnoses and the delivery of high-quality patient care, as it helps clinicians navigate between various diagnostic options and adjust their thinking for both simple and complex cases¹.

Clinical reasoning is built on three main pillars: framing the encounter (Table 1), diagnostic reasoning (Table 2), and managing patient care (Table 3). These pillars help clinicians accurately identify active medical problems, identify diagnoses and risk factors, and develop patient-appropriate treatments. Large Language Models (LLMs) have recently emerged as a tool to supplement and enhance clinicians’ reasoning. In this Comment, we discuss the potential of LLMs to provide clinician-level clinical reasoning using these pillars as a guideline.

Clinical reasoning

The first pillar of clinical reasoning, framing the clinical encounter, centers around gathering and organizing patient information to pinpoint active issues, evaluate severity, and lay the groundwork for further analysis (Table 1). LLMs can mimic this process through their ability to extract, categorize, and synthesize data from both structured and unstructured clinical sources². With the appropriate prompts, an LLM can assess a patient’s electronic health record, including notes, lab results, and symptoms, to create a structured overview that highlights key clinical findings. This approach reflects the initial phases of clinician reasoning, where data is considered based on prior knowledge and contextual significance². LLMs can categorize symptoms, deduce timelines, and prioritize important findings, providing an initial triage assessment.

However, due to the lack of direct patient interaction, LLMs depend significantly on the clarity and completeness of the information received. The strength of an LLM lies in organizing existing data rather than noticing subtleties such as facial expressions, tone of voice, or bedside demeanor—areas where human touch is irreplaceable. Consequently, while an LLM can

help in synthesizing information, it still requires a clinician to apply judgment and contextual understanding^{2,3}.

Diagnostic reasoning, the second pillar, entails clinicians developing, refining, and prioritizing a list of potential diagnoses utilizing both probabilistic insights and their own experiential knowledge (Table 2). LLMs tackle this task by aligning patient information with vast training datasets, which can encompass textbooks, clinical protocols, and de-identified case histories. An LLM can mimic reasoning by pinpointing statistically important correlations between observed symptoms and potential diagnoses, relying heavily on large-scale pattern recognition^{2,3}. When given a clinical vignette, an LLM can suggest a ranked list of differential diagnoses, complete with justifications based on relevant features. This process reflects how clinicians mentally retrieve illness scripts, using cognitive frameworks that connect risk factors, underlying mechanisms, and clinical manifestations¹. In addition, LLMs can be guided to evaluate and differentiate diagnostic options using distinguishing characteristics, akin to the approach of seasoned clinicians.

However, LLMs do not perform hypothesis testing or engage in metacognitive analysis. They fail to assess internal consistency or adjust based on new information unless prompted again. For instance, a recent study involving AI-assisted virtual urgent care indicated that diagnoses and management plans generated by LLMs aligned with final decisions made by clinicians in over half of the cases, often receiving higher ratings for optimality than those proposed by clinicians^{3,4}. Nonetheless, clinicians excelled in adapting to new or evolving information, a key trait of expert clinical judgment. Therefore, while LLMs contribute to and enhance diagnostic reasoning through pattern recognition and data retrieval, human oversight is essential for contextual adaptability and error identification.

The third pillar, treatment and management, focuses on prioritizing interventions, customizing therapy according to individual patient factors, and monitoring responses over time (Table 3). LLMs can be beneficial in this process by retrieving and summarizing evidence-based treatment guidelines, recommending both pharmacologic and non-pharmacologic options, and offering comparative effectiveness information on various therapeutic choices³. A recent randomized controlled trial revealed that doctors utilizing GPT-4 assistance performed significantly better in management reasoning tasks than clinicians using conventional resources⁵. It is also important to note that LLMs spent more time per case, and there was no significant difference between LLM-augmented clinicians and LLMs alone⁵. This finding suggests that large language models can improve clinical decision-making in intricate cases⁵. With a confirmed diagnosis, an LLM can delineate standard care management plans, detailing first-line treatments, alternative options for patients with contraindications, and supplementary

¹University of Rochester, Rochester, NY, USA. ²Cardiovascular Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.

³Broad Institute, Massachusetts Institute of Technology, Cambridge, MA, USA. e-mail: armoundas.antonis@mgh.harvard.edu

Table 1 | Strategic integration across pillar 1 of clinical reasoning: framing the clinical encounter (information gathering and organization)

LLM Utility		LLMs can mimic the initial process of gathering and organizing patient information by extracting, categorizing, and synthesizing data from both structured and unstructured clinical sources like EHRs, notes, lab results, and symptoms, by creating a structured overview that highlights key clinical findings, effectively providing an initial triage assessment.
Clinician Imperatives	Direct Patient Engagement	Conducting of thorough patient interview and physical examination to gather subjective information, observe non-verbal cues, and build rapport—elements that LLMs cannot provide due to their lack of direct patient interaction.
	Contextual Validation	Verifying the LLM-generated summary against the patient's lived experience and current context. Information from an EHR might be outdated or incomplete, and the LLM will not inherently know this without explicit prompting.
	Prompt Engineering	Mastering the art of crafting precise and comprehensive prompts to guide the LLM in extracting the most relevant information for a specific clinical scenario.

Table 2 | Strategic integration across pillar 2 of clinical reasoning: diagnostic reasoning (developing and refining diagnoses)

LLM Utility		LLMs can significantly assist in generating and refining differential diagnoses by aligning patient information with vast training datasets of medical knowledge, including textbooks, clinical guidelines and de-identified case histories. They can suggest a ranked list of potential diagnoses with justifications, reflecting how clinicians mentally retrieve illness scripts.
Clinician Imperatives	Critical Evaluation and Refinement	Clinicians must critically evaluate every diagnostic suggestion from an LLM. This involves considering the likelihood of each diagnosis, its consistency with all available patient data, and ruling out alternatives through targeted questioning, physical examination, and diagnostic testing.
	Considering Nuance and Ambiguity	LLMs struggle with ambiguity, as their design prioritizes pattern matching over navigating uncertainty in complex, real-world clinical scenarios. They may produce convincing but inaccurate diagnoses with incomplete or misleading data, highlighting the importance of human oversight. Human clinicians excel in adapting to new or evolving information, a key trait of expert clinical judgment.
	Addressing Atypical Presentations and Rare Conditions	LLMs, being pattern-matching systems, may overlook rare diseases or atypical presentations that fall outside their learned patterns. Human intuition and experience are crucial for considering the full spectrum of possibilities.
	Cognitive Bias Awareness	While LLMs have their own biases, human clinicians must also be aware of their own cognitive biases (e.g., anchoring) and actively work to mitigate them.

Table 3 | Strategic integration across pillar 3 of clinical reasoning: treatment and management (prioritizing interventions and customizing care)

LLM Utility		LLMs can retrieve and summarize evidence-based treatment guidelines, recommending treatment options, and offering comparative effectiveness information. They can delineate standard care management plans, detailing first-line treatments, alternative options, and supplementary measures. LLMs can also create easily understandable patient education materials.
Clinician Imperatives	Individualized Patient Care	Treatment decisions must be individualized, considering patient comorbidities, allergies, current medications, lifestyle, socioeconomic factors, and personal preferences—elements an LLM cannot fully grasp without explicit, detailed, and often unquantifiable input.
	Monitoring and Adjustment	Clinicians are responsible for continuously modifying their approaches based on patient interactions and monitoring responses over time. LLMs do not inherently perform this iterative process unless continuously fed new data and re-prompted.
	Ethical and Economic Considerations	While LLMs could potentially query insurance or cost implications (depending on access to sensitive patient data), clinicians must remain the primary arbiters of ethical considerations, resource allocation, and ensuring equitable access to care. Decisions about costly treatments should always prioritize patient well-being over purely economic efficiency.
	Patient Education and Shared Decision-Making	While LLMs can generate educational materials, the clinician's role in addressing patients' questions and concerns, conveying complex medical information empathetically, and facilitating true shared decision-making is paramount.

measures to enhance symptom management⁵. Moreover, LLMs can support shared decision-making by creating easily understandable patient education materials that clarify the risks, benefits, and expected outcomes of different treatments. Since, treatments can be costly, LLMs could help in querying a patient's insurance, depending on the type of LLM, and the access it has to patient information.

Despite these capabilities, LLMs work in a fundamentally different manner than human clinicians. LLMs depend solely on established correlations from their training data³. This means that, although an LLM can produce thorough diagnoses, it will lack genuine clinical intuition, which is

the ability to detect when a presentation strays from typical cases. Additionally, the inability to conduct independent hypothesis testing means an LLM does not “think” like humans do; it will not question assumptions, acknowledge cognitive biases, or critically evaluate the reliability of incoming information. LLMs are also challenged by ambiguity, as the design of an LLM prioritizes pattern matching over navigating uncertainty in complex, real-world clinical scenarios². Consequently, they may produce convincing but inaccurate diagnoses when faced with incomplete or misleading data, highlighting the importance of human oversight to refine and validate diagnostic hypotheses.

Delineating LLM capabilities from human cognition

Considering the three main pillars of clinical reasoning, it appears that LLMs are powerful statistical engines that excel at identifying patterns and relationships within vast datasets, but they lack consciousness, intuition, or genuine clinical judgment. Thus, a foundational principle to be observed by clinicians engaging with LLMs is a clear understanding of the capabilities and, crucially, the limitations.

The primary strength of an LLM lies in the ability to rapidly process and synthesize immense volumes of textual data, including medical literature, clinical protocols, and de-identified patient histories. This allows them to extract, categorize, and organize information from complex sources, identifying statistical correlations that underpin diagnostic suggestions or treatment recommendations, reducing the time spent on data review and information gathering, and allowing clinicians to focus on higher-level tasks.

In contrast, human clinicians bring a multifaceted set of skills that LLMs cannot replicate. These involve: direct patient interaction and observation, which includes the ability to observe subtle non-verbal cues (e.g., facial expressions, tone of voice), interpret complex emotional states, and engage in empathetic communication, which are critical for building rapport and understanding the full patient narrative. LLMs also lack contextual understanding, which is the capacity to integrate medical knowledge with the unique psychosocial context of each patient, including their values, preferences, socioeconomic factors, and cultural background. This nuanced understanding is often not explicitly captured in structured data but is vital for personalized care. LLMs also lack clinical intuition and expertise which is developed through years of experience. Clinicians intuitively recognize patterns and can determine when a presentation deviates from typical cases or when to consider rarer conditions. While LLMs rely on large-scale pattern recognition from training data, they lack this genuine intuition. People can also engage in metacognitive analysis and hypothesis testing, questioning their assumptions, recognizing cognitive biases (e.g., anchoring) and critically evaluating the reliability of incoming information, and engaging in iterative hypothesis testing as new data emerges.

Therefore, the paradigm should always include “a human in-the-loop”. LLMs are sophisticated tools to *augment* human capabilities, not to replace them. Clinicians should strategically integrate LLMs into each pillar of clinical reasoning, leveraging strengths while vigilantly addressing limitations, ultimately maintaining the responsibility for all clinical decisions and patient care outcomes (Tables 1–3).

Bias and ethical considerations affecting clinical reasoning

It is known that there is bias across artificial intelligence (AI) algorithms^{6,7} and LLMs. This is because models reflect biased information present in their training datasets.

Thus, clinicians must be acutely aware that LLM outputs may perpetuate historical biases in medical data, potentially leading to suboptimal or inequitable care for certain demographic groups^{8,9}. This can manifest as underrepresentation in training samples, leading to predictive performance disparities. Every LLM-generated recommendation should be scrutinized for any signs of bias related to race, gender, socioeconomic status, or other protected attributes^{8,9}. If a recommendation seems incongruous or suggests a differential treatment approach based on non-clinical factors, it should be thoroughly questioned and cross-referenced with unbiased evidence. Consequently, clinicians should advocate for and support the development of LLMs trained on diverse, representative, and carefully curated datasets to minimize inherent biases.

Beyond bias, clinicians must consider broader ethical implications. While currently the responsibility remains with the clinicians when an LLM augments and aids decision-making, rather than replacing it, the question of accountability when an algorithmic decision causes unintended harm becomes critical as LLMs streamline decision-making processes^{6,10,11}. Clinicians should strive for transparency with patients about the use of AI tools in their care, explaining how LLMs assist in decision-making and emphasizing continued human oversight^{6,10,11}; the goal should be “ante-hoc

interpretability,” where the model’s soundness and human intelligibility are guaranteed by design, rather than “post-hoc explainability”, which can be misleading. Furthermore, it must be ensured that the use of LLMs does not diminish patient agency⁴ by allowing LLMs to shift authority away from humans.

Upholding data privacy and security standards

The integration of LLMs into healthcare systems necessitates stringent adherence to data privacy and security regulations, given the sensitive nature of patient health information¹². Clinicians should understand how patient data is handled, stored, and used by the LLM provider. If using LLMs for research or non-direct patient care, all patient data should be properly de-identified before input.

Individual clinical practice must be supported by broad systemic safeguards. Clinicians have a crucial role in advocating for robust regulatory and ethical frameworks. They should emphasize the need for rigorous, independent clinical validation studies for LLMs before widespread adoption, especially for high-stakes clinical decisions. The current lack of large-scale, prospective, clinical trials investigating patient outcomes from LLM use is a significant gap^{6,10,11}.

Clinicians should support the development of clear legal and ethical accountability frameworks for LLM-driven errors or harms in healthcare and contribute to the development of professional guidelines and best practices for the responsible use of LLMs within medical societies and regulatory bodies.

Finally, clinicians should advocate for AI systems that seamlessly integrate into and *augment* well-established medical workflows and real-life reasoning processes, as opposed to disrupting or fully automating them. This “machine-in-the-loop” approach acknowledges that AI should complement human abilities, boost effectiveness, and champion clinical best practice, rather than attempt to “solve” medical challenges algorithmically and potentially limit human exploration or progress.

Conclusions

LLMs constitute a powerful technological advance with the potential to significantly augment clinical reasoning and improve patient care. However, the utility of LLMs is maximized when they are approached as a class of sophisticated tools that require informed oversight, critical evaluation, and a deep understanding of their inherent capabilities and limitations, rather than replacements for human judgment. In this Comment, we have discussed the potential of LLMs to contribute to clinical reasoning. Some models show impressive accuracy in diagnostic reasoning and management, particularly with diagnostic cases, while others struggle to outperform specialized multidisciplinary clinical teams⁸.

The inherent limitations stated above point to the ongoing requirement for human oversight and the development of strong frameworks to ensure effective collaboration between AI systems and healthcare providers. By adhering to the proposed recommendations, clinicians can harness the power of LLMs safely and effectively, ultimately advancing the quality and efficiency of healthcare delivery while preserving the essential human element of medicine. Creating strategies for AI-human collaboration will be crucial for harnessing LLMs’ capabilities while upholding high standards of patient care and safety.

Received: 4 July 2025; Accepted: 20 October 2025;

Published online: 20 November 2025

References

1. Bowen, J. L. Educational strategies to promote clinical diagnostic reasoning. *N. Engl. J. Med.* **355**, 2217–2225 (2006).
2. Gu, B., Desai, R. J., Lin, K. J. & Yang, J. Probabilistic medical predictions of large language models. *NPJ Digital Med.* **7**, 367 (2024).
3. Brodeur, P. G. et al. Superhuman performance of a large language model on the reasoning tasks of a physician. *arXiv*. <https://arxiv.org/abs/2412.10849Oxford>.

4. Armondas, A. A. & Loscalzo, J. Patient agency and large language models in worldwide encoding of equity. *NPJ Digital Med.* **8**, 258 (2025).
5. Goh, E. et al. GPT-4 assistance for improvement of physician performance on patient care tasks: a randomized controlled trial. *Nat. Med.* **31**, 1233–1238 (2025).
6. Armondas, A. A. et al. Use of Artificial Intelligence in Improving Outcomes in Heart Disease: A Scientific Statement From the American Heart Association. *Circulation*. 2024. **In an effort to address the complex issues pertaining to the transformation healthcare delivery with artificial intelligence (AI), the American Heart Association produced its first Scientific Statement on AI and cardiovascular disease and care, aiming to present the state-of-the-art on best practices, gaps and challenges in improving the applicability of AI algorithms in each domain of cardiovascular disease and care.**
7. Sevakula, R. K. et al. State-of-the-art machine learning techniques aiming to improve patient outcomes pertaining to the cardiovascular system. *J. Am. Heart Assoc.* **9**, e013924 (2020).
8. Bazoukis, G. et al. Impact of social determinants of health on cardiovascular disease. *J. Am. Heart Assoc.* **14**, e039031 (2025).
9. Pearson, T. A. et al. The science of precision prevention: Research opportunities and clinical applications to reduce cardiovascular health disparities. *JACC Adv.* **3**, 100759 (2024).
10. Bazoukis, G. et al. The inclusion of augmented intelligence in medicine: A framework for successful implementation. *Cell Rep. Med.* **3**, 100485 (2022).
11. Bota, P., Thambiraj, G., Bollepalli, S. C. & Armondas, A. A. Artificial intelligence algorithms in cardiovascular medicine: An attainable promise to improve patient outcomes or an inaccessible investment? *Curr. Cardiol. Rep.* **26**, 1477–1485 (2024).
12. Spector-Bagdady, K. et al. Principles for health information collection, sharing, and use: A policy statement from the American Heart Association. *Circulation* **148**, 1061–1069 (2023).

Acknowledgements

A.A.A. is funded, in part, by the Institute of Precision Medicine (17UNPG33840017) of the American Heart Association, the RICBAC Foundation, NIH grants 1 R01 HL135335-01, 1 R01 HL161008-01, 1 R21 HL137870-01, 1 R21EB026164-01 and 3R21EB026164-02S1.

Author contributions

M.C.: Participated in the writing and reviewing of the manuscript. A.A.A.: Participated in the conception of the study, the writing and reviewing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Antonis A. Armondas.

Peer review information *Communications Medicine* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at

<http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025