NEJM AI

NEJM AI 2026;3(11)
DOI: 10.1056/AIp2600773

PERSPECTIVE

No Mastery without the Machine — Medical Education in the Age of Clinical AI

Philip Jarrett, M.D., M.B.A.,[1] D. Mark Courtney, M.D., M.Sci.,[1] Deborah Diercks, M.D., M.S., M.B.A.,[1] and Samia Farooqi, M.D.[1]

Received: June 6, 2026; Revised: August 26, 2026; Accepted: September 9, 2026; Published: October 13, 2026

Abstract

AI is often framed as a threat to medical education owing to concerns it may diminish the cognitive struggle through which trainees traditionally develop clinical judgment — a legitimate but incomplete concern. Medicine has never been defined by unaided human cognition; it has been defined by the obligation to use the best available tools in the service of patients. As AI systems increasingly extend not only clinicians' perception, but also their reasoning, medical mastery may shift from memorization and independent synthesis toward the orchestration, supervision, and calibrated deference to intelligent systems. The task for medical education is therefore not to preserve cognitive struggle for its own sake, but to refocus it from recalling facts to interrogating machine reasoning, detecting failure modes, and knowing when human intervention improves care versus when it merely interrupts a better system.

Introduction

For more than a century, medical education has treated cognitive struggle as a necessary element of professional formation. Students memorize biochemical pathways, defend differential diagnoses, present patients on rounds, and learn through the discomfort of being uncertain in front of others. This struggle is not incidental. It is one way trainees learn judgment, humility, and the limits of pattern recognition.

The arrival of generative AI has therefore prompted understandable concern. If trainees can ask a model to generate a differential diagnosis, critique a management plan, summarize a record, or draft a note, will they still develop the cognitive muscle required to practice medicine? Will a generation trained with AI become clinically fluent but intellectually hollow?[1]

These are serious concerns. But they risk mistaking the preservation of a training method for the preservation of medical mastery itself.

AI and the Changing Nature of Clinical Expertise

Medicine has never been defined by unaided human cognition, but by the obligation to use the best available tools in the service of patients. The stethoscope extended hearing. The electrocardiogram extended perception. The laboratory extended inference. Imaging

The author affiliation is listed at the end of the article.

Philip Jarrett can be contacted at philip.jarrett@utsouthwestern.edu or at Department of Emergency Medicine, UT Southwestern Medical Center, 5323 Harry Hines Blvd E4.300, Dallas, TX 75390.

NEJM AI is produced by NEJM Group, a division of the Massachusetts Medical Society.
Downloaded from ai.nejm.org by EMANUELA TERESA LOCATI on October 22, 2026.

extended anatomy beyond the visible body. The electronic health record, despite its many failures, extended memory. Each technology displaced some prior form of clinical skill while also giving clinicians new skills to hone.

Artificial intelligence may be different in degree, but not in kind. It does not merely extend the physician's senses; it increasingly extends the physician's reasoning. That is precisely why it feels threatening. AI enters the territory physicians have long regarded as most central to professional identity: the capacity to synthesize incomplete information, generate hypotheses, estimate risk, and recommend action.[2]

Yet the relevant question is not whether AI will erode some traditional forms of expertise. It will. The question is whether medical education will defend yesterday's expertise as a marker of professional legitimacy or cultivate the expertise required for tomorrow's care.

The analogy may be less chess than aviation. We do not train pilots to navigate primarily by stars because celestial navigation once represented mastery. We train them to understand aircraft systems, weather, automation, communication, failure modes, and the conditions under which human intervention is lifesaving or catastrophic. The modern pilot is not less skilled because autopilot exists. The locus of skill has shifted from continuous manual control to systems command.

Medicine may be entering a similar transition. The physician of the near future may not be the person who independently recalls the longest differential diagnosis or personally integrates every clinical variable. Instead, the physician may be the person who can orchestrate intelligent systems: directing AI agents to obtain history, interrogate the medical record, identify inconsistencies, generate diagnostic and therapeutic pathways, surface uncertainty, communicate trade-offs, and monitor for deterioration (Fig. 1). This is not "never skilling."[3] It is a different skill.

Refocusing the Struggle

This reframing should not become an excuse for frictionless training. A trainee who has never struggled to form an argument, defend a differential diagnosis, or sit with uncertainty may be poorly prepared to recognize when an AI system is incomplete, overconfident, or wrong. The danger is not that trainees will use AI. The danger is that they may never learn what it feels like to reason before AI intervenes.

The task, then, is not to remove struggle from medical education. It is to refocus it. Future trainees may need less friction around memorizing facts that machines can retrieve instantly and more friction around interrogating machine reasoning, reconciling contradictory outputs, identifying hidden assumptions, detecting automation bias, and deciding when deference to AI is safer than human correction. The educational objective should not be unaided cognition for its own sake, but calibrated judgment in the presence of powerful cognitive machinery.

Implications for Medical Education and Practice

This shift also exposes the limitations of current assessment and supervision. A curriculum organized around closed-book recall and standardized testing may be poorly suited to an era in which the central competency is not memory alone, but combined human and machine judgment. Training will require faculty who can model nondefensive AI use at the bedside and beyond, AI systems that serve as tutors as well as tools, and assessments that test whether physicians can recognize unsupported recommendations, explain algorithmic reasoning, and justify accepting or overriding a machine-generated plan.

The evidence base is still incomplete, but the trajectory is difficult to ignore. Recent studies suggest that frontier language models can match or exceed physicians on some complex diagnostic and management reasoning tasks.[4] Other work has shown that physician use of AI is not automatically beneficial; the human–AI team can perform worse when clinicians misunderstand, underuse, or overtrust model output.[5] In the field of surgery, AI-enabled robotic systems have begun performing complex procedural tasks in controlled settings with increasing autonomy.[6] None of this proves that AI is ready to replace clinicians in ordinary practice. Current systems remain brittle, variably evaluated, unevenly integrated into workflows, and vulnerable to failure modes that may be invisible to patients and clinicians alike.

But dismissing these systems because they are not yet complete is like dismissing the Wright Flyer because it could not cross the Atlantic. The relevant fact is not that early aircraft were fragile; it is that flight had become possible.

The most comforting arguments against medical AI are also the most precarious. We say AI cannot understand the physician–patient relationship. We say it cannot do what

NEJM AI is produced by NEJM Group, a division of the Massachusetts Medical Society.
Downloaded from ai.nejm.org by EMANUELA TERESA LOCATI on October 22, 2026.

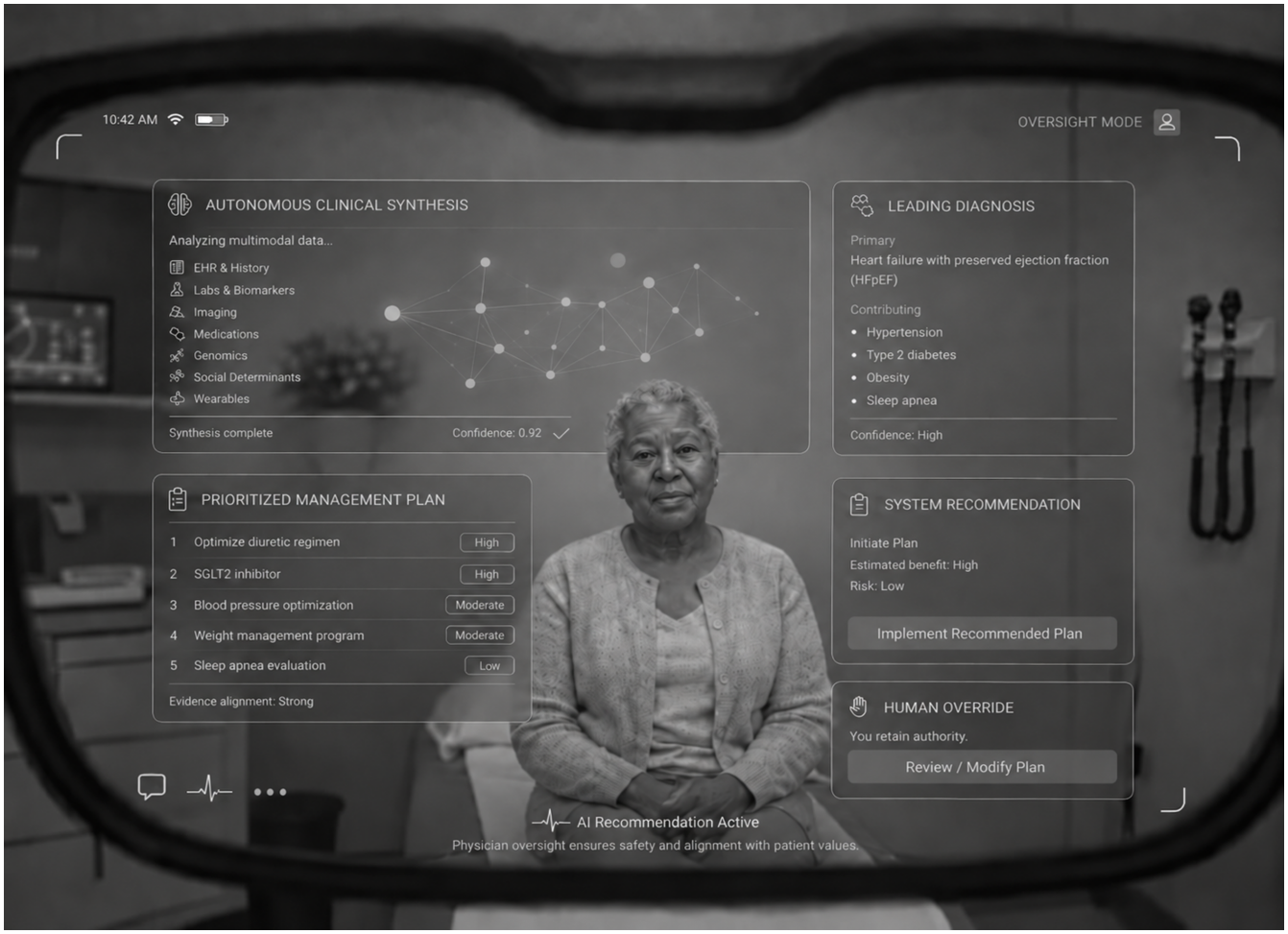


Figure 1. Physician Judgment in an AI-Mediated Clinical Encounter.

This figure shows a hypothetical clinical interface in which AI generates diagnostic and management recommendations while the physician evaluates their appropriateness, incorporates patient priorities, and judges when to defer or intervene. Displayed findings and confidence estimates are illustrative. EHR denotes electronic health record; and SGLT2, sodium–glucose cotransporter-2.

we do. We say patients need human presence, moral judgment, and empathy. In many clinical contexts, these claims remain true. But they can also become professional defense mechanisms — ways of converting current indispensability into a permanent moral principle.

The physician–patient relationship is sacred because it serves the patient, not because it preserves the physician. If an AI system someday diagnoses more accurately, treats more safely, monitors more continuously, communicates more patiently, and remembers more completely than a human clinician, our ethical obligation will not be to protect the primacy of the physician; it will be to ask what role the physician should continue to occupy.

This possibility should not be caricatured as technological surrender. It is a call for professional humility. Medical education should not teach students to compete with AI on tasks machines may soon perform better. It should teach them to evaluate, supervise, challenge, integrate, and, when necessary, yield to AI.

Yielding to AI is undoubtedly an uncomfortable dynamic to accept. Physicians are trained to act, decide, rescue, and take responsibility. We are less often trained to recognize when our intervention adds noise to a better system. Yet the future of medicine may require a new form of restraint: knowing when human judgment improves care and when it merely interrupts it.

NEJM AI is produced by NEJM Group, a division of the Massachusetts Medical Society.
Downloaded from ai.nejm.org by EMANUELA TERESA LOCATI on October 22, 2026.

The Future Role of the Physician

For centuries, physicians have stood at the center of medicine's cognitive architecture. AI asks whether that position is a law of nature or a historical phase.[2] The profession can respond by insisting that authentic medical practice requires unaided human mastery. Or it can prepare trainees for a world in which mastery means knowing how to command systems more capable than oneself.

The first path protects the mythology of the physician. The second protects patients.

One day, the most dangerous actor in the room may not be the machine making a recommendation; it may be the human who insists on overruling it simply because being in the loop still feels like being in charge.

Disclosures

Author disclosures are available at ai.nejm.org.

ChatGPT Images 2.5 (OpenAI) was used to assist with image generation. No AI tools were used for writing assistance.

Author Affiliation

[1]Department of Emergency Medicine, The University of Texas Southwestern Medical Center, Dallas, USA

References

1. Keren R, Desai BR, West DC. Promoting clinical expertise in the age of AI: no struggle, no mastery. JAMA 2026;335:1837-1838. DOI: 10.1001/jama.2026.6097.
2. Brodeur PG, Buckley TA, Kanjee Z, et al. Performance of a large language model on the reasoning tasks of a physician. Science 2026;392:524-527. DOI: 10.1126/science.adz4433.
3. Ke Y, Jin L, Chiat Ling Ong J, et al. AI-induced never-skilling in medical education. Nat Med 2026;32:1997–2006. DOI: 10.1038/s41591-026-04438-y.
4. Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Netw Open 2024;7:e2440969. DOI: 10.1001/jamanetworkopen.2024.40969.
5. Rodman A, Beam AL, Manrai AK. Exploring the past — and future — of medical diagnosis and artificial intelligence. NEJM AI 2024;1(9). DOI: 10.1056/AIp2400707.
6. Kim JWB, Chen JT, Hansen P, et al. SRT-H: a hierarchical framework for autonomous surgery via language-conditioned imitation learning. Sci Robot 2025;10:eadt5254. DOI: 10.1126/scirobotics.adt5254.

NEJM AI is produced by NEJM Group, a division of the Massachusetts Medical Society.
Downloaded from ai.nejm.org by EMANUELA TERESA LOCATI on October 22, 2026.