

Reconsidering trust at the bedside in the era of artificial intelligence

Vijay Rajput, MD, MACP, FRCP



Dr. Rajput is Professor and Chair of the Department of Medical Education at the Nova Southeastern University, Dr. Kiran C. Patel College of Allopathic Medicine in Fort Lauderdale, FL. His email address is: vrajput@nova.edu.

Introduction

On my first day of internship in the United States, I believed I spoke “good English” and understood medicine well. I was wrong on both counts. When my first patient told me he had “passed out,” I was alarmed—where I trained, that phrase meant death. When he added that he had “thrown up,” I was lost again. Even the nurse’s comment about having a “bad hair day” left me confused.¹ I quickly learned that medicine is not only practiced through language, knowledge, and protocols, but through trust, culture, context, and human connection. What helped me survive and later thrive was not superior diagnostic reasoning but learning how patients recognize whether a physician genuinely cares about them as a human being and can be trusted.

More than three decades later, medicine stands at another inflection point. Artificial intelligence now assists with diagnoses, predictions, documentation, and even empathic responses. Some studies suggest that, in text-based scenarios, AI chatbots are rated as more empathic than human clinicians. Several mechanisms may explain this paradox. AI-generated responses are linguistically calibrated, consistently structured, and free from time pressure or emotional fatigue. Unlike clinicians operating in high-cognitive-load environments, chatbots provide uninterrupted, validation-rich language and allow patients time to process information asynchronously. In this sense, perceived empathy may reflect not deeper understanding, but the optimization of communication conditions rarely available in real-world clinical encounters.² A recent systematic review and meta-analysis found that AI chatbot responses were rated significantly higher in empathy than those of clinicians, particularly in text-only environments lacking nonverbal cues.² That finding is both fascinating and unsettling. If machines can simulate empathy convincingly, what becomes of the human physician at the bedside? The answer lies in understanding what trust really is—and what it is not.

Reconsider trust

In emotionally charged or ethically complex situations, patients may also perceive AI systems as more “neutral” and less judgmental, potentially reducing defensive emotional responses that can arise in human-to-human interactions. However, such neutrality may come at the cost of moral engagement, as trust in medicine depends not only on the absence of judgment, but on the presence of humility and accountable care.

Trust is not efficiency. It is not accuracy alone. And it is not a well-phrased response generated in milliseconds. Trust, at its core, is a willingness to be vulnerable in the presence of another human being. It is based on the expectation that the other person is competent, benevolent, and acts with integrity. This classic framework, long applied in organizational settings, maps remarkably well onto the patient–physician relationship.³ AI can support competence. It may even enhance it. But benevolence and integrity are perceived and analyzed, not computed. Recent research shows that when AI-based clinical decision support systems are introduced into care, patient trust can subtly erode, particularly in high-risk situations and among certain populations. In these settings, patients may perceive physicians as less competent or less personally invested, even when AI recommendations are sound.⁴

The issue is not that AI is harmful; it is that its presence can unintentionally distance clinicians from patients if not carefully mediated. Emergency department studies reinforce this lesson. Patient trust is strongly associated with communication, perceived empathy, transparency, and the feeling that expectations are being met more often than with technical outcomes themselves. Long waits and uncertainty erode trust, but so does the sense that no one is truly listening or caring. These dynamics are further compounded by systemic pressures—time constraints, administrative burden, and clinician burnout—which can inadvertently degrade communication quality and relational presence at the bedside. This is where humanism becomes not a nostalgic ideal, but a clinical necessity.⁵

Trust at the bedside is developed by small, often invisible acts: sitting instead of standing, acknowledging fear, explaining uncertainty, and inviting patients into decisions that affect their lives. Physicians who become patients often develop a heightened awareness of vulnerability, communication gaps, and system inefficiencies. This lived experience frequently reshapes their

approach to care, reinforcing that empathy is not merely expressed, but deeply informed by having “been there.”

As emphasized by Abraham Verghese, MD (AQA, East Tennessee State University Quillen College of Medicine, 1989, Faculty), the “laying on of hands” represents more than physical examination—it is a symbolic and relational act that affirms presence, attention, and care. Such embodied practices remain central to the experience of being seen and valued as a patient.⁶ The shared decision-making, when done authentically, has been shown to strengthen trust particularly by improving communication and respecting patient values. These are domains that may remain difficult to fully delegate to algorithms. And yet, there is a paradox worth confronting. In controlled, text-only environments, AI responses are often perceived as more empathic than those written by clinicians (chatbot versus human).² This should not embarrass us. It should instruct us. The problem is not that machines are becoming more human. The problem is that humans, under time pressure, burnout, and system constraints, are being forced to communicate in increasingly transactional ways.

AI exposes our deficits by reflecting them back to us. The danger is not that AI will replace physicians, but that it will normalize a version of care where empathy is scripted, trust is assumed, and presence is optional. When patients sense indifference, whether from a human distracted by a screen or a machine delivering polished reassurance, trust is likely to erode. Once lost, trust is remarkably difficult to restore. Restoring trust requires intentional humanistic practice. It requires teaching trainees that professionalism is not just about competence, but about how patients feel in their presence. It requires reminding clinicians that no algorithm can substitute for the moral act of showing up fully for another person. Empirical and theoretical studies note that trust is foundational to patient–physician interactions and that failure to recognize patients’ personal concerns and preferences can erode trust and bedside humanism.⁷

Agentic AI and the new fragility of bedside trust

Beyond decision-support tools, medicine is now entering the era of agentic AI—systems capable of autonomously initiating actions, setting goals, prioritizing tasks, and interacting with patients and clinicians with minimal human oversight. Examples include AI agents that triage messages, generate care plans, adjust orders, or

communicate recommendations directly to patients. While these systems promise efficiency and scalability, they introduce a qualitatively different threat to bedside trust.

Trust in medicine depends not only on outcomes, but on accountability. Patients implicitly ask: Who is responsible for this decision? Agentic AI complicates that question. When clinical actions are initiated or shaped by autonomous systems, responsibility becomes diffused across clinicians, institutions, and opaque algorithms. Research in human–AI interaction shows that perceived loss of clinician agency can undermine trust, particularly when patients sense that decisions are being “handed off” to systems they do not understand or consent to engage with them.^{8,9}

Agentic AI may also erode trust by altering relational dynamics. Trust is built through deliberation, explanation, and moral presence—processes that unfold in dialogue. Autonomous systems optimize speed and task completion, not for relational meaning. Studies in healthcare automation demonstrate that higher levels of autonomy are associated with reduced perceptions of empathy, diminished transparency, and lower patient confidence, even when technical performance improves.⁴ When patients encounter care that feels predetermined or nonnegotiable, they may interpret efficiency as indifference.

Moreover, agentic AI risks normalizing moral distancing. Clinicians may become supervisors of outputs rather than authors of decisions and outcome, weakening patients’ perception of benevolence and integrity—core components of trust. As ethicists have warned, delegating judgment to machines can erode the moral practice of medicine itself.¹⁰

Agentic AI therefore demands explicit boundaries. Autonomy must be constrained, accountability and/or responsibility must remain human, and patients must experience clinicians as fully present moral agents. Without these safeguards, the very systems designed to streamline care may quietly corrode the trust upon which medicine depends. This moment also demands reflection within the profession. Trust at the bedside is not a birthright; it is earned through consistent demonstration of competence, communication, and care. If AI exposes gaps in how we listen, respond, and connect, then the appropriate response is not defensiveness, but renewal. Medicine must “heal itself” by recommitting to relational excellence alongside technical progress.

Conclusion

AI should free us to be more human, not less. Used wisely, AI can reduce cognitive load, surface hidden risks, and create space for clinicians to do what only humans can do: interpret suffering, navigate uncertainty, and bear witness. But this requires vigilance. We must resist the temptation to outsource not just thinking, but caring. I learned early in my training that patients quickly recognize whether a physician genuinely cares and is trustworthy. That lesson remains unchanged even as technology evolves. Amid artificial intelligence, the future of medicine may not be secured by smarter machines alone, but by clinicians who remember that trust is built one human encounter at a time. It is also important to acknowledge uncertainty. AI-enabled systems may, in some contexts, outperform average clinicians in diagnostic accuracy, consistency, and even perceived empathy. If so, patients and health systems may gravitate toward models that optimize output at scale. The question, therefore, is not whether physicians will remain indispensable, but what aspects of care cannot be reduced to optimization alone, particularly those involving meaning, trust, and moral accountability.

References

1. Rajput V. Not born in the USA. *JAMA*. 2012;308(21):2197–2198. doi:10.1001/jama.2012.45065.
2. Howcroft A, Bennett-Weston A, Khan A, Griffiths J, Gay S, Howick J. AI chatbots versus human healthcare professionals: A systematic review and meta-analysis of empathy in patient care. *British Medical Bulletin*. 2025;156(1), ldaf017. <https://doi.org/10.1093/bmb/ldaf017>.
3. Mayer RD, JH; Schoorman, FD. An integrative model of organizational trust. *Academy of Management Review*. 1995;20(3):709–34.
4. Zondag AGM, Rozestraten R, Grimmelikhuijsen SG, Jongsma KR, van Solinge WW, Bots ML, Vernooij RWM, Haitjema S. The effect of artificial intelligence on patient-physician trust: Cross-sectional vignette study. *J Med Internet Res*. 2024;26:e50853.
5. Abidova A, Alcantara da Silva P, Moreira S. The role of confidence/trust in the emergency department. *BMC Research Notes*. 2025;18:198. <https://doi.org/10.1186/s13104-025-07266-4>.
6. Verghese A. Treat the patient, not the CT scan. *N Engl J Med*. 2011;364(25):2471–3. doi:10.1056/NEJMp1101548 doi: 10.2196/50853 PMID: 38805702 PMCID: 11167322
7. Birkhäuser J, Gaab J, Kossowsky J, Hasler S, Krummenacher P, Werner C, et al. Trust in the health care professional and health outcome: A meta-analysis. *PLoS ONE*. 2017;12(2): e0170988. <https://doi.org/10.1371/journal.pone.0170988>.
8. Esmailzadeh P, Mirzaei T, Dharanikota S. Patients' perceptions toward human-artificial intelligence interaction in health care: Experimental study. *J Med Internet Res*. 2021;23(11):e25856. doi: 10.2196/25856. PMID: 34842535; PMCID: PMC8663518.
9. Longoni C, Bonezzi A, Morewedge CK. . Resistance to medical artificial intelligence. *Journal of Consumer Research*. 2019;46(4), 629–50. <https://doi.org/10.1093/jcr/ucz013>
10. Nagy M, Jones CB. How will artificial intelligence affect patient-clinician relationships? *AMA J Ethics*. 2020;22(5):E395-400. doi: 10.1001/amajethics.2020.395. <https://journalofethics.ama-assn.org/article/how-will-artificial-intelligence-affect-patient-clinician-relationships/2020-05>