Journal of Studies in Medical Sciences; Vol. 1, No. 1, 2025
ISSN 2810-9899
Published by July Press

Microsoft AI Diagnostic Orchestrator (MAI-DxO) Promises Cost Savings in Diagnosis — But at What Cost to the Medical Profession?

Malik Sallam[1,2], Chadia Beaini[3,4], Maad M. Mijwil[5,6] & Mohammed Sallam[4,7,8,9]

[1] Department of Pathology, Microbiology and Forensic Medicine, School of Medicine, The University of Jordan, Amman, Jordan

[2] Department of Clinical Laboratories and Forensic Medicine, Jordan University Hospital, Amman, Jordan

[3] Department of Nephrology/Dialysis, Mediclinic Welcare Hospital, Mediclinic Middle East, Dubai, United Arab Emirates

[4] College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences (MBRU), Dubai, United Arab Emirates

[5] College of Administration and Economics, Al-Iraqia University, Baghdad, Iraq

[6] Computer Techniques Engineering Department, Baghdad College of Economic Sciences University, Baghdad, Iraq

[7] Department of Pharmacy, Mediclinic Parkview Hospital, Mediclinic Middle East, Dubai, United Arab Emirates

[8] Department of Management, Mediclinic Parkview Hospital, Mediclinic Middle East, Dubai, United Arab Emirates

[9] Department of Management, School of Business, International American University, Los Angeles, CA, USA

Correspondence: Malik Sallam, Department of Pathology, Microbiology and Forensic Medicine, School of Medicine, The University of Jordan, Amman 11942, Jordan. Tel: 962-79-184-5186. E-mail: malik.sallam@ju.edu.jo

Received: June 2, 2025 Accepted: July 7, 2025 Online Published: July 9, 2025

doi:10.20849/jsms.v1i1.1515 URL: https://doi.org/10.20849/jsms.v1i1.1515

Abstract

The deployment of superintelligent generative artificial intelligence (genAI) in clinical diagnosis represents one of the most consequential transformations in the history of medicine. Advanced tools such as the Microsoft AI Diagnostic Orchestrator (MAI-DxO), in collaboration with OpenAI's Generative Pre-trained Transformer (GPT)-derived models, now outperform generalist physicians in diagnostic accuracy and efficiency, boasting accuracy rates exceeding 85% while reducing testing costs by up to 70%. Advocates hail these breakthroughs as a triumph of scale, speed, and economic rationality—offering hope to under-resourced healthcare systems especially in low- and middle-income countries (LMICs) and overwhelmed clinicians. Yet this economic promise masks a deeper, more insidious cost; the creeping marginalization of the physician's role. Framed as partners in care, clinicians increasingly face relegation to supervisory or validation functions, as algorithms undertake the core cognitive tasks that have long defined the medical profession. The physician risks being transformed from healer and thinker to technician and observer. If left unchecked, this trajectory portends the de-skilling of medical practice, the erosion of professional identity, and the eventual displacement of human clinicians in favor of algorithmic workflows. This *Perspective* examines the global economic logic driving the adoption of superintelligent genAI in diagnosis, critiques the ethical and professional consequences of unchecked automation, and proposes guardrails to protect the centrality of human clinical judgment. Building on recent findings—such as the MAI-DxO diagnostic benchmark, even though it has not yet been approved for clinical use—and informed by historical examples of professional disruption, we argue that the rise of superintelligent genAI in medicine demands active, deliberate oversight rather than passive acceptance. As with past technological revolutions, the outcome will not be decided by capability alone, but by values. If cost-saving becomes the singular metric of success, we may soon find ourselves in a healthcare system where efficiency thrives—but care, in its fullest human sense, does not.

Keywords: generative AI, ethical governance of AI, physician displacement, job loss, medical professionalism

1. Introduction

The latest extraordinary development in diagnostic medicine has not come from a university hospital or a Nobel laureate. It came from a consortium of engineers, researchers, and physicians working in tandem with Microsoft and OpenAI, who released a tool they believe might achieve what no individual physician ever could; mastery over all of medicine. On June 30, 2025, Microsoft revealed what it called the path to medical superintelligence; the Microsoft AI Diagnostic Orchestrator (MAI-DxO) (King et al., 2025). When paired with state−of−the−art generative AI (genAI) models—most notably OpenAI's o3—it achieved diagnostic accuracy of 85–85.5% on 304 *New England Journal of Medicine* clinical cases, compared to around 20% for 21 experienced physicians working unaided (Limb, 2025; Nori et al., 2025). Specifically, MAI-DxO did more than achieve higher diagnostic accuracy; it ordered fewer unnecessary tests, cutting diagnostic costs by up to 70% compared to off−the−shelf models, and 20% compared to physicians (Nori et al., 2025).

The aforementioned metrics may sound utopian, but their appeal is grounded in hard realities. From the perspectives of public health and fiscal responsibility, the promise is undeniable. In the United States, nearly 30% of total healthcare spending—part of a system that is estimated to consume approximately 20% of the gross domestic product (GDP)—can be considered as a waste, including unnecessary testing, avoidable hospitalizations, and administrative inefficiencies (Berwick & Hackbarth, 2012; Shrank et al., 2019). Thus, the specter of artificial intelligence (AI)−driven diagnostics reducing such waste appeals equally to policymakers and patients (Khanna et al., 2022; Singh & Kaunert, 2025). Additionally, in low- and middle-income countries (LMICs) settings, where physician density might be insufficient and diagnostic delay is deadly, the prospect of ubiquitous, expert−level AI diagnosis is nothing short of revolutionary (Ciecierski-Holmes et al., 2022; Khan et al., 2024; Dangi et al., 2025).

Yet there is a darker—less discussed—side to this AI transformation in healthcare (Sallam et al., 2025b). History teaches us that no technology is born simply to augment; it also threatens to supersede. Consider the telegraph's displacement of the horse−courier, or the printing press' sidelining of scribes. It is unlikely that MAI-DxO and its future superintelligent companions will remain a silent partner forever. When a machine can match or surpass human performance at lower cost, it becomes a substitute, not a supplement. However, Microsoft, for its part, maintains that MAI-DxO will complement clinicians by stating that "*Clinical roles are much broader than diagnosing. They require human trust, judgment, and empathy*"—qualities AI lacks (King et al., 2025). Microsoft AI Chief Executive Officer (CEO), speaks of a near future of medical superintelligence, but emphasizes that the AI is not yet ready for clinical deployment (Milmo, 2025). Meanwhile, researchers acknowledge the AI technology's power in healthcare but also its limits; machines may diagnose, but they cannot care (Sallam, 2023; Booven et al., 2025; Rawat et al., 2025; Takita et al., 2025). The clinical diagnosis is not only pattern recognition—it is bearing witness to suffering, interpreting ambiguity, holding guilt and hope in equal measure (Elstein & Schwartz, 2002; Pelaccia et al., 2015; Croskerry et al., 2023). AI may calculate probabilities, but it cannot narrate context, weigh values, negotiate trade−offs, or sit vigil beside a dying patient (Gönd öcs & D örfler, 2024; Rubin et al., 2025).

More ominously, MAI-DxO's developers show no intention of slowing down. Microsoft plans to embed the technology across its products—from Bing health searches to Copilot assistants—and integrate it into global care pathways (King et al., 2025). As AI performance improves, healthcare cost shrinks, and regulatory oversight lags, the economic forces driving automation will accelerate (Bajwa et al., 2021). Soon, the discourse will shift from whether AI can diagnose more accurately than physicians to whether physicians should be diagnosing at all (Rodler et al., 2024; Fransen et al., 2025; Guo & Chen, 2025; Lenharo, 2025). If we pause now to analyze, we see two forces pulling at the healthcare seams. On one side is efficiency; the sharpened edge of performance and cost−control that promises leaner, faster, more accurate care (Krishnan et al., 2023; Jeanty et al., 2024; Taherdoost & Ghofrani, 2024; McDuff et al., 2025). On the other side there is a growing—and justified—apprehension among current and future healthcare professionals; that the very foundations of the medical profession are at risk (Rony et al., 2024; Malik Sallam et al., 2024; Elfaham et al., 2025; Sallam et al., 2025a). Medicine is not solely a technical pursuit; it is built on the cumulative traditions of knowledge, empathy, judgment, and professional ethos that have defined the physician's role for centuries (Hannawa et al., 2022). If healthcare is to endure with its integrity and humanity intact, the promise of AI technological efficiency must be reconciled with the preservation of these core professional values (Mennella et al., 2024).

Thus, this *Perspective* aimed to examine the potential unintended consequences of introducing superintelligent genAI diagnostic tools into clinical medicine—specifically, their impact on the role, identity, and future of the physician. Beyond questions of performance and efficiency lie deeper concerns; the erosion of clinical judgment, the displacement of professional expertise, and the social consequences of outsourcing diagnostic authority to

machines. What may begin as technological augmentation risks becoming professional obsolescence. The clinician, once a central figure in diagnostic reasoning, risks being reduced to a validator of machine output. Medical education may shift away from cultivating independent judgment and clinical curiosity toward training clinicians to navigate and confirm algorithmic AI decisions. Over time, the diagnostic wisdom that has defined medicine for centuries may atrophy—not from lack of relevance, but from lack of use. This *Perspective* aimed to confront that possibility head-on and propose a path forward that preserves the physician's essential role in the age of superintelligent AI.

2. Superintelligent AI in Clinical Diagnosis: The Erosion of Clinical Judgment and Diagnostic Authority

In the tradition of medicine, diagnosis has never been a rote act of classification. It has been, instead, a form of inquiry—a process that unites science with narrative, statistics with intuition, and pattern recognition with patient individuality (Rosenberg, 2002; Jutel, 2009; Holmboe & Durning, 2014). The physicians do not merely catalog symptoms; they interpret meaning. They listen not only to what is said, but also to what is withheld (Sarafis et al., 2013; Varkey, 2021). Diagnosis, in its highest form, is less about identifying pathology than about understanding the person in whom it resides (Rosenberg, 2002).

And yet, we now stand at the threshold of a profound healthcare practice transformation brought about by AI. With the advent of superintelligent AI diagnostic tools such as MAI-DxO—models capable of outperforming generalist physicians across hundreds of complex cases—diagnosis could soon be algorithmic (King et al., 2025; Nori et al., 2025). These superintelligent AI models lean to map input to output with statistical precision. And they do so with competence so overwhelming that it threatens to displace the very foundation of human clinical reasoning.

Consider a real-life analogy; the skilled pilot and the autopilot. When automated systems first entered the cockpit, they were designed to assist. But as the technology improved, pilot skills atrophied. Incidents such as the 2009 crash of Air France Flight 447 tragically demonstrated that when systems failed, human pilots—long detached from the manual controls—might be ill-prepared to swiftly intervene (Brown, 2016; Oliver et al., 2017). In medicine, we risk a similar fate. If AI becomes the default diagnostic tool, the physician may one day find themselves in the position of the pilot who forgot how to fly (Cnossen, 2014; Quinn et al., 2021; George et al., 2024).

Already, this AI shift is visible across various medical specialties (Shafi & Parwani, 2023; Brady et al., 2024; Reddy, 2024; Mart ńez-Vargas et al., 2025). This trend is reinforced by recent comparative studies of genAI performance compared to humans (Takita et al., 2025). For example, in one large, multi-national investigation, a GPT-4-turbo–powered virtual assistant was tested against 17,144 physicians from Italy, France, Spain, and Portugal, each responding to medical licensing exam questions in their native language (Miranda et al., 2025). The genAI consistently outperformed physicians across all countries and knowledge domains, achieving accuracy rates between 72% and 96%, compared to 46% to 62% among human counterparts (Miranda et al., 2025). Particularly notable was the finding that performance declined with physician seniority—suggesting that AI not only surpasses the average clinician but may outperform even experience itself (M. Sallam et al., 2024; Miranda et al., 2025). Only in pediatrics did physicians match or modestly exceed the genAI, and even then, the margin was narrow and statistically insignificant (Miranda et al., 2025).

The clinical implications of superintelligent diagnostic AI are far from benign. When AI-based tools consistently generate accurate differential diagnoses and associated management plans, the physician's role in diagnostic reasoning risks quiet erosion. As these AI tools demonstrate consistent success, the incentive to engage in independent, critical analysis weakens—not from negligence, but from habituation. Over time, clinicians may shift from diagnosticians to overseers of pre-processed outputs, with diminishing inclination to question the machine's conclusions. This gradual shift is not merely theoretical; it reflects a broader transformation in clinical identity, training, and authority. Key themes in this diagnostic reconfiguration are summarized in (Table 1).

Table 1. Themes on the erosion of clinical judgment in the age of superintelligent artificial intelligence (AI)

Theme	Summary
Rise of superintelligent AI	Systems like MAI-DxO outperform physicians in accuracy, tempting healthcare to substitute, not support, human reasoning.
The automation paradox	Like the autopilot in aviation, over-reliance on AI risks skill atrophy. Physicians may forget how to "fly" independently.
Educational consequences	Medical students increasingly rely on AI tools, risking underdevelopment of diagnostic thinking akin to using calculators for arithmetic.
De-skilling and role shift	Physicians risk being reduced to software supervisors, replacing independent judgment with algorithm validation.

MAI-DxO: Microsoft AI Diagnostic Orchestrator

Medical education, too, is bending to this AI-induced shift (Kundu, 2021). Students are being taught not to wrestle with complexity, but to consult clinical decision support tools (Zhai et al., 2024). A generation raised on AI-enhanced platforms may excel in identifying what the machine missed—but at the cost of their own diagnostic development (Gerlich, 2025). It is akin to learning arithmetic by calculator: the output is reliable, but the mental architecture behind the numbers is never built (George et al., 2024). This evolution may be inevitable, but it is not without consequence. The physician, once an originator of insight, risks becoming a system operator. Their role devolves from thoughtful interpretation to the supervision of software. Clinical judgment, cultivated through years of mentorship, bedside experience, and epistemic humility, would be flattened into interface fluency and algorithmic oversight; a feared outcome that should be mitigated by human empathy and ethical oversight (Çetinkaya, 2025). Moreover, the shift from thinking diagnostician to diagnostic validator introduces a moral hazard: the de-skilling of the medical profession (Abd-Alrazaq et al., 2023; Arif et al., 2023; Knopp et al., 2023; Preiksaitis & Rose, 2023; Safranek et al., 2023; Sallam et al., 2023; Tsang, 2023; Abdullah et al., 2024; Janumpally et al., 2024; Li & Li, 2024; Xu et al., 2024; Mondal, 2025; Natali et al., 2025).

In addition, the burden of accountability in the context of AI-assisted clinical diagnosis remains unresolved (Habli et al., 2020; Nichol et al., 2023). When physicians are no longer expected to formulate differentials themselves, but merely to approve or sign off on those generated by a machine, their moral and legal responsibility becomes ambiguous (Durán & Jongsma, 2021; Nichol et al., 2023). In such a milieu, responsibility becomes diffuse—no longer clearly attributable to either the clinician or the developers of the AI (Choudhury & Asan, 2022; Bouderhem, 2024). This diffusion of accountability poses a serious threat to patient safety, professional integrity, and public trust (Habli et al., 2020; Harishbhai Tilala et al., 2024). However, some may argue that such change is progress—that it allows clinicians to focus on empathy and communication while the machine handles the technical reasoning (Sallam, 2023; Maleki Varnosfaderani & Forouzanfar, 2024). This division between technical reasoning and empathetic engagement is, in reality, a false dichotomy. In clinical practice, empathy is not a separate or secondary task; it is embedded within the diagnostic process itself. The physician's ability to recognize delayed care-seeking, to interpret the subjective framing of symptoms, or to question atypical presentations is not ancillary to diagnosis—it is diagnostic (Harada et al., 2024). These interpretive functions rely on contextual awareness, narrative competence, and relational insight, which might not be replicable by algorithmic AI systems (Bjerring & Busch, 2021; Sun et al., 2024).

To dissociate diagnostic reasoning from its human context also risks optimizing accuracy while impoverishing the therapeutic relationship. The central risk posed by superintelligent genAI in clinical care is not its failure, but its success. As genAI tools achieve and eventually surpass human diagnostic accuracy, there is a growing risk that the practice of medicine will be restructured around their capabilities (Fogo et al., 2024). The interpretive art of diagnosis—once characterized by curiosity, ethical reflection, and critical dialogue—may be displaced by computational expediency (Al-kfairy et al., 2024). The physician's role could be reduced to reviewing automated outputs, responding to dashboards, or confirming suggested plans, rather than exercising independent clinical judgment. This AI revolution would mark a profound shift in the epistemology of medicine, replacing open-ended reasoning with probabilistic prediction, and substituting dialogic engagement with protocol adherence (Schmidt et al., 2025). The physician's greatest asset—the capacity to think, to question, and to contextualize uncertainty—must not be supplanted by a model of care in which human deliberation is deemed inefficient or redundant. As genAI advances, the challenge is not to halt its integration into clinical workflows, but to ensure that its adoption does not come at the cost of the professional identity, moral responsibility, and

intellectual agency that have long defined the practice of medicine (Hassan et al., 2024; Yim et al., 2024).

3. Superintelligent AI Ethical Risks: Accountability, Equity, and Loss of Human Presence

As the clinical integration of genAI accelerates, we must not lose sight of the ethical terrain it threatens to reshape as shown in (Figure 1). Beyond questions of diagnostic performance or professional displacement lie deeper risks—those of accountability, equity, and the erosion of human presence in medicine. These risks are not speculative; they are already materializing at the clinical, legal, and societal institutions (Tigard, 2021; Naik et al., 2022; Wang et al., 2023).

Figure 1. Ethical Risks of Superintelligent AI in Clinical Medicine

First, there is the urgent matter of AI accountability (Ayari & Ammar, 2024). When a diagnostic AI tool errs, the chain of responsibility becomes murky (Coeckelbergh, 2023). If a physician overrides the AI suggestion and patient harm results, they may be held personally liable. Conversely, if a clinician follows the AI recommendation and harm still occurs, can they reasonably be expected to defend a decision they did not generate, did not fully understand, and were perhaps subtly compelled to accept by institutional protocol or legal risk. In either case, the physician bears the moral burden, while the developer or deploying entity of the AI tool remains shielded behind layers of technical abstraction and regulatory ambiguity (Shumway & Hartman, 2024; Krügel et al., 2025). The current medico-legal framework is yet to be fully-equipped to address this diffusion of responsibility in AI assisted decisions (Bleher & Braun, 2022; Bartlett, 2023).

Unlike traditional tools—stethoscopes, laboratory tests, or even imaging modalities—AI tools are autonomous, non-transparent, and dynamically evolving (Fehr et al., 2024). Their "reasoning" is not interpretable in the human sense, and their outputs are probabilistic, not categorical. This raises critical questions for malpractice litigation, informed consent, and patient autonomy (Gibelli et al., 2024). For example, a patient cannot meaningfully consent to a diagnosis derived from an AI algorithm they cannot interrogate, especially when the physician overseeing the process may not fully understand the model's internal logic. Until legal standards evolve in parallel with technological deployment, we risk burdening clinicians with liability for decisions they did not truly make (Hurley et al., 2025).

Second, there is the risk of AI exacerbating an already existent inequity. GenAI models, including MAI-DxO, are trained predominantly on structured, high-quality clinical datasets—drawn disproportionately from academic centers, insured populations, and cases with rich documentation (Kaushal et al., 2020; Ferrara, 2024; Wei et al., 2025). These data sources, while excellent for algorithmic training, are not representative of the full diversity of

human illness, particularly in marginalized communities where care may be sporadic, records incomplete, or symptom narratives non-standard (Allareddy et al., 2023; Currie et al., 2025). This imbalance is not benign. AI tools trained on such data will predict well for those who resemble the dataset—and poorly for those who do not (Cross et al., 2024). Racial and ethnic minorities, rural populations, those with rare or poorly understood conditions, and patients with complex social determinants of health may be further marginalized by diagnostic engines optimized for a more "legible" population (Mittermaier et al., 2023; Cau et al., 2025). In effect, we risk building a two-tier system of medical precision; one where the majority receives increasingly accurate diagnoses, while the vulnerable receive approximations—elegantly wrong rather than contextually right. This is not mere theory since early and recent studies have shown that AI algorithms used in radiology underperformed in detecting pathologies among women, racial and ethnic minorities, younger individuals, and patients of lower socioeconomic status (Mamary et al., 2018; Seyyed-Kalantari et al., 2021). As AI becomes a diagnostic gatekeeper, it risks encoding historical biases into future care unless equity is proactively engineered into its development, deployment, and regulation (Chen et al., 2024; Cross et al., 2024).

Finally—and perhaps most insidiously—there is the erosion of the human presence in healthcare. Diagnosis is not merely the identification of pathology. It is a moment of shared vulnerability, a dialogue through which the patient begins to make sense of their suffering. A physician does not merely say "You have cancer." They sit with the patient in that moment, explain its implications, anticipate their fears, and begin—often silently—the long work of psychological adaptation (Tranberg et al., 2024). This is the dimension of care that cannot be rendered into code. No matter how sophisticated or superintelligent an AI tool becomes, it will not ask, "How are you holding up?" It will not detect that a tremor in the voice matters more than the content of the words. It will not recognize when silence is not compliance, but despair. In the words of William Osler, "*the good physician treats the disease; the great physician treats the patient who has the disease*" (Gibson, 2015).

To displace the physician from this moment is to risk hollowing out medicine from the inside. We may achieve diagnostic perfection—but at the cost of relational bankruptcy. While superintelligent genAI tools may deliver accurate diagnostic labels and treatment plans, they do not provide meaning, context, or guidance for living with illness. Precision alone does not constitute care. The risk is that, in prioritizing correct answers, we reduce medicine to a technical exercise—efficient but impersonal. Integrating AI tools like MAI-DxO into healthcare practice requires more than performance optimization; it demands transparency, clinical oversight, and attention to equity (Sallam, 2023; Singh & Keche, 2025). Even the most accurate diagnosis is only one element of what patients seek. They turn to clinicians not only for answers, but for understanding, presence, and human connection.

4. Toward Ethical Superintelligent AI Integration in Healthcare: Recommendations for Governance

In every age of technological upheaval, civilization is presented with a choice: to let innovation dictate our values or to ensure our values guide innovation. The integration of genAI into healthcare—particularly in clinical diagnosis—presents such a moment. Like the printing press in the 15th century or electrification in the 20th, genAI promises unprecedented gains in efficiency and reach. But without deliberate governance, AI also threatens to displace professional wisdom, erode patient trust, and depersonalize care. Thus, as genAI evolves rapidly across the various healthcare settings, there is an urgent need for collaborative action among all stakeholders—clinicians, technologists, policymakers, ethicists, and patients—to develop clear, enforceable guidelines that ensure ethical, transparent, and responsible use of these tools (Sallam, 2024). To preserve the moral and intellectual integrity of medicine while embracing its technological future, a framework of active, principled AI governance is essential (Jha et al., 2025). In this *Perspective*, we propose three interdependent domains for such governance: clinical oversight, economic alignment, and medical education. Each is a bulwark against the commodification of care, and each demands urgent, coordinated action as shown in (Figure 2).

Figure 2. Ethical Governance Pillars for Integrating Superintelligent Generative Artificial Intelligence (genAI) in Clinical Diagnosis

4.1 Preserve Physician Oversight as a Standard of Care

Regardless of the sophistication of superintelligent genAI, the final authority in clinical decision-making must remain vested in the physician. This is not a matter of professional protectionism, but one of ethical necessity, epistemic prudence, and public trust. The regulatory infrastructure that governs clinical practice—including licensing boards, accrediting agencies, and health system policies—must formally codify this principle; AI may inform, but it must not autonomously determine diagnostic or therapeutic decisions (Mennella et al., 2024). In this framework, superintelligent genAI should be conceptualized not as an autonomous agent, but as a high-functioning diagnostic adjunct—comparable to the role of imaging or laboratory results. Just as a radiograph requires expert interpretation to inform patient care, so too must AI-generated outputs be subject to human analysis, contextual integration, and professional judgment. To allow otherwise would constitute a category error—confusing computational output with clinical reasoning.

The risk of removing human oversight is not merely theoretical. Historical parallels illustrate the danger of replacing judgment with automation in complex, value-laden domains. Consider the analogy of the judiciary; replacing a trial judge with an algorithm that produces higher rates of correct verdicts may improve outcomes statistically—but it would also erode the legitimacy of legal proceedings (Selçık et al., 2025). Process, interpretation, and moral accountability are integral to institutional trust. Medicine, like law, is not simply about results; it is about how those results are reached, by whom, and under what ethical commitments. Furthermore, the need for physician oversight becomes especially urgent in cases involving overlapping symptomatology, psychosocial complexity, or diagnostic ambiguity—scenarios in which the art of medicine is most critical. In such contexts, genAI tools lack the capacity to perceive nonverbal cues, detect inconsistencies in patient narratives, or spontaneously ask follow-up questions that are not explicitly prompted. They cannot infer distress from hesitation, nor can they say, "Something doesn't feel right." These capacities are central to the physician's interpretive role and are irreducible to computational processes.

Crucially, patient trust is anchored in the belief that medical decisions are made by a clinician who not only possesses technical expertise but also bears moral responsibility for the outcomes. If patients come to believe that their diagnoses are determined by automated systems rather than human caregivers who understand their histories, values, and concerns, the physician-patient relationship may suffer irreparable harm—even in the presence of accurate outcomes. Therefore, preserving physician oversight is not merely a best practice; it must be established as a non-negotiable standard of care. Just as we do not permit laboratory results to dictate

management without clinical correlation, neither should we permit AI-generated diagnoses to bypass human adjudication. Clinical judgment must remain the final arbiter—not despite AI's power, but because of it.

4.2 Align Financial Incentives With Professional Roles

Health systems are shaped not only by ethical aspirations but by economic architecture. Regardless of professional ideals, the structure of reimbursement exerts a profound influence on clinical practice. As genAI enters diagnostic workflows, the risk is not merely technical—it is economic. The prevailing temptation is to deploy these systems as substitutes for human labor, justified by projected gains in efficiency, throughput, and cost-containment. Left unchecked, this logic may gradually displace physician-led decision-making and diminish the value of clinical judgment itself. The consequences of such displacement extend beyond workflow optimization. When the diagnostic process is delegated to algorithms operating without meaningful human oversight, the profession risks erosion of its cognitive identity. In such a model, physicians are no longer diagnosticians, but validators of machine output—an arrangement that undermines both patient safety and the long-standing ethical compact between physician and society.

To prevent this trajectory, payment models must be realigned to affirm the indispensable role of the clinician. Specifically, reimbursement frameworks should explicitly differentiate between autonomous AI pathways and those in which a licensed physician evaluates, contextualizes, and, where appropriate, modifies the AI's recommendations. Financial recognition should be assigned not merely to the act of diagnosis, but to the interpretive labor involved in exercising clinical oversight. This approach would reflect the added value of human judgment—particularly in complex, ambiguous, or ethically sensitive cases where AI performance may be limited or biased. By contrast, diagnostic processes that rely exclusively on AI without clinician engagement should not be incentivized through cost savings alone. Such cases should trigger additional layers of clinical review, regulatory scrutiny, and outcome auditing. The goal is not to impede innovation but to ensure that the implementation of AI respects the professional structures that secure safe, equitable, and accountable care.

In short, economic incentives must not undermine the physician's central role in diagnosis. A system that rewards automation over clinical participation risks substituting statistical performance for humane care. While a machine may indeed be faster or cheaper, it cannot substitute the moral, relational, and contextual dimensions that define healing. Reimbursement strategies must reflect this fundamental distinction, reinforcing the principle that technological advancement should serve—not replace—professional expertise.

4.3 Reimagine Medical Education for an AI-Augmented Future

Preparing the next generation of clinicians for a future shaped by superintelligent diagnostic systems requires a thoughtful recalibration of medical education. This evolution should not entail the abandonment of traditional pedagogies, but rather their deliberate expansion. As AI becomes increasingly embedded in clinical practice, medical training must equip physicians with the skills to critically appraise AI-generated outputs, identify limitations in model inference, and judiciously integrate algorithmic recommendations into complex, context-dependent decision-making as highlighted by several recent studies (Hopson et al., 2024; Schubert et al., 2025; Schuitmaker et al., 2025; Thompson et al., 2025).

Technical fluency with AI must be considered necessary but not sufficient. The fundamental pillars of medical training—diagnostic reasoning, ethical reflection, and narrative competence—must not be supplanted by interface navigation or algorithmic literacy. Indeed, the clinician of the AI era must not be trained as a technician with an M.D. degree, but as a physician whose identity is rooted in cognitive autonomy and professional leadership (Siala & Wang, 2022). Physicians should be trained to collaborate with AI tools while maintaining the capacity to assume full diagnostic control when the system's output is ambiguous, biased, or clinically inappropriate (Moreno-Sánchez et al., 2025). To that end, medical curricula should incorporate formal instruction in areas such as probabilistic reasoning, interpretability of machine learning (ML) models, and human-AI interaction (Abd-Alrazaq et al., 2023; Pregowska & Perkins, 2024). This should include case-based simulations in which trainees are asked to evaluate AI recommendations within the broader clinical context, weighing patient-specific factors, psychosocial determinants, and ethical considerations (Abdullah et al., 2024).

However, these additions must not displace the core humanistic competencies of medicine. The ability to synthesize narrative, question assumptions, and appreciate ambiguity must remain central. The patient is not a vector of data points to be optimized, but a human being whose illness exists within a social, psychological, and moral framework. Ultimately, the aim of AI-integrated medical education should not be to generate passive users of diagnostic tools, but active stewards of clinical judgment in an age of intelligent systems. Physicians must not cede epistemic authority to the machine, but rather, engage it critically—wielding its power without surrendering their own. In doing so, we safeguard not the intellectual integrity of the medical profession and the moral and

relational dimensions of healthcare that no algorithm, however advanced, can replace.

5. Conclusions

The rise of superintelligent genAI in clinical diagnosis marks a turning point in the history of medicine. The performance gains are real. The economic efficiencies are persuasive. But if implemented without deliberation, the superintelligent genAI systems risk transforming medicine from a healing profession into a predictive enterprise; precise, cost-effective, yet profoundly dehumanized. The physician is not merely a conduit for clinical information. They are the moral and cognitive anchor of the diagnostic encounter. At least in the short term, no machine, however advanced, can replicate the act of listening, the burden of deciding, or the grace of presence. In our pursuit of superintelligent machines, we must not create a system that no longer needs the physician—only to discover too late that it no longer cares for the patient either. This *Perspective* calls for integration of superintelligent AI into medicine with wisdom, humility, and foresight—not to replace ourselves, but to enhance what only humans can offer.

Finally, certain limitations of this *Perspective* must be acknowledged. The current work was derived from secondary data, conceptual frameworks, and historical analogies, rather than from direct empirical evidence gathered through prospective, real-world clinical studies. Its primary focus remains on diagnostic genAI tools; thus, therapeutic applications, administrative automation, and other dimensions of genAI lie beyond its present scope. Meaningful progress in this field will require rigorous, longitudinal investigations assessing the actual clinical, economic, and humanistic impacts of tools such as MAI-DxO across varied healthcare settings. Particular emphasis must be placed on patient outcomes, physician professional well-being, and the distribution of benefits and risks across diverse populations. Furthermore, advancing this field demands a parallel effort to develop coherent legal, ethical, and educational frameworks that govern the responsible deployment of superintelligent AI—ensuring alignment with the foundational principles of medicine. The lack of such empirical inquiry and systemic safeguards risks that healthcare policy would remain locked in a reactive posture—perpetually trailing technological advances rather than shaping their trajectory. The future of medicine cannot hinge on technological capability alone; it must be steered by deliberate, evidence-based governance that safeguards both human dignity and clinical excellence.

Abbreviations

AI: Artificial intelligence

CEO: Chief Executive Officer

GDP: Gross domestic product

genAI: Generative artificial intelligence

GPT: Generative pre-trained transformer

LMICs: Low- and middle-income countries

MAI-DxO: Microsoft AI Diagnostic Orchestrator

ML: Machine learning

Author Contributions

Conceptualization, Malik Sallam; methodology, Malik Sallam, Chadia Beaini, Maad M. Mijwil, Mohammed Sallam; software, Malik Sallam; investigation, Malik Sallam, Chadia Beaini, Maad M. Mijwil, Mohammed Sallam; resources, Malik Sallam; data curation, Malik Sallam, Chadia Beaini, Maad M. Mijwil, Mohammed Sallam; writing—original draft preparation, Malik Sallam; writing—review and editing, Malik Sallam, Chadia Beaini, Maad M. Mijwil, Mohammed Sallam; visualization, Malik Sallam; supervision, Malik Sallam; project administration, Malik Sallam. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Not applicable.

Acknowledgments

This study used ChatGPT-4o for language refinement (improving grammar, sentence structure, and readability of the manuscript). We confirm that all AI-assisted processes were critically reviewed by the authors to ensure the integrity and reliability of the results. The final decisions and interpretations presented in this article were solely made by the authors.

Conflicts of Interest

The authors declare no conflicts of interest.

References

Abd-Alrazaq, A., AlSaad, R., Alhuwail, D., *et al.*. (2023). Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions. *JMIR Med Educ*, *9*, e48291. https://doi.org/10.2196/48291

Abdullah, H. M. A., Naeem, N. I., Malkana, G. A., *et al.*. (2024). Exploring the Ethical Implications of ChatGPT in Medical Education: Privacy, Accuracy, and Professional Integrity in a Cross-Sectional Survey. *Cureus*, *16*(12), e75895. https://doi.org/10.7759/cureus.75895

Al-kfairy, M., Mustafa, D., Kshetri, N., *et al.*. (2024). Ethical Challenges and Solutions of Generative AI: An Interdisciplinary Perspective. *Informatics*, *11*(3), 58. https://doi.org/10.3390/informatics11030058

Allareddy, V., Oubaidin, M., Rampa, S., *et al.*. (2023). Call for algorithmic fairness to mitigate amplification of racial biases in artificial intelligence models used in orthodontics and craniofacial health. *Orthod Craniofac Res, 26 Suppl 1*, 124-130. https://doi.org/10.1111/ocr.12721

Arif, T. B., Munaf, U., & Ul-Haque, I. (2023). The future of medical education and research: Is ChatGPT a blessing or blight in disguise?. *Med Educ Online*, *28*(1), 2181052. https://doi.org/10.1080/10872981.2023.2181052

Ayari, A., & Ammar, A. (2024). AI Accountability, Ethics, and Human Resource Implications. In A. Ammar, & M. Bernstein (Eds.), *Ethical Challenges for the Future of Neurosurgery* (pp. 47-56). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-71477-1_4

Bajwa, J., Munir, U., Nori, A., *et al.*. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthc J*, *8*(2), e188-e194. https://doi.org/10.7861/fhj.2021-0095

Bartlett, B. (2023). The possibility of AI-induced medical manslaughter: Unexplainable decisions, epistemic vices, and a new dimension of moral luck. *Medical Law International*, *23*(3), 241-270. https://doi.org/10.1177/09685332231193944

Berwick, D. M., & Hackbarth, A. D. (2012). Eliminating waste in US health care. *Jama*, *307*(14), 1513-1516. https://doi.org/10.1001/jama.2012.362

Bjerring, J. C., & Busch, J. (2021). Artificial Intelligence and Patient-Centered Decision-Making. *Philosophy & Technology*, *34*(2), 349-371. https://doi.org/10.1007/s13347-019-00391-6

Bleher, H., & Braun, M. (2022). Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems. *AI Ethics*, *2*(4), 747-761. https://doi.org/10.1007/s43681-022-00135-x

Booven, D. V., Cheng-Bang, C., & Meenakshy, M. (2025). Chapter 8 - Limitations of artificial intelligence in healthcare. In H. Arora (Ed.), *Artificial Intelligence in Urologic Malignancies* (pp. 231-246). Academic Press. https://doi.org/B978-0-443-15504-8.00008-9

Bouderhem, R. (2024). Shaping the future of AI in healthcare through ethics and governance. *Humanities and Social Sciences Communications*, *11*(1), 416. https://doi.org/10.1057/s41599-024-02894-w

Brady, A. P., Allen, B., Chong, J., *et al.*. (2024). Developing, purchasing, implementing and monitoring AI tools in radiology: practical considerations. A multi-society statement from the ACR, CAR, ESR, RANZCR & RSNA. *Insights into Imaging*, *15*(1), 16. https://doi.org/10.1186/s13244-023-01541-3

Brown, J. (2016). The Effect of Automation on Human Factors in Aviation. *Journal of Instrumentation, Automation and Systems*, *3*, 31-46. https://doi.org/10.21535/jias.v3i2.916

Cau, R., Pisu, F., Suri, J. S., *et al.*. (2025). Addressing hidden risks: Systematic review of artificial intelligence biases across racial and ethnic groups in cardiovascular diseases. *European Journal of Radiology*, *183*, 111867. https://doi.org/10.1016/j.ejrad.2024.111867

Çetinkaya, L. (2025). Redefining Mentorship in Medical Education with Artificial Intelligence: A Delphi Study on the Feasibility and Implications. *Teach Learn Med*, 1-11. https://doi.org/10.1080/10401334.2025.2521001

Chen, F., Wang, L., Hong, J., *et al.*. (2024). Unmasking bias in artificial intelligence: a systematic review of bias detection and mitigation strategies in electronic health record-based models. *J Am Med Inform Assoc*, *31*(5), 1172-1183. https://doi.org/10.1093/jamia/ocae060

Choudhury, A., & Asan, O. (2022). Impact of accountability, training, and human factors on the use of artificial intelligence in healthcare: Exploring the perceptions of healthcare practitioners in the US. *Human Factors in Healthcare*, *2*, 100021. https://doi.org/10.1016/j.hfh.2022.100021

Ciecierski-Holmes, T., Singh, R., Axt, M., *et al.*. (2022). Artificial intelligence for strengthening healthcare systems in low- and middle-income countries: a systematic scoping review. *NPJ Digit Med*, *5*(1), 162. https://doi.org/10.1038/s41746-022-00700-y

Cnossen, F. (2014). Cognitive Skill in Medicine: An Introduction. *PanVascular Medicine*, 1-41. https://doi.org/10.1007/978-3-642-37393-0_181-1

Coeckelbergh, M. (2023). Narrative responsibility and artificial intelligence. *AI & SOCIETY*, *38*(6), 2437-2450. https://doi.org/10.1007/s00146-021-01375-x

Croskerry, P., Campbell, S. G., & Petrie, D. A. (2023). The challenge of cognitive science for medical diagnosis. *Cogn Res Princ Implic*, *8*(1), 13. https://doi.org/10.1186/s41235-022-00460-z

Cross, J. L., Choma, M. A., & Onofrey, J. A. (2024). Bias in medical AI: Implications for clinical decision-making. *PLOS Digit Health*, *3*(11), e0000651. https://doi.org/10.1371/journal.pdig.0000651

Currie, G. M., Hawk, K. E., & Rohren, E. M. (2025). Generative Artificial Intelligence Biases, Limitations and Risks in Nuclear Medicine: An Argument for Appropriate Use Framework and Recommendations. *Semin Nucl Med*, *55*(3), 423-436. https://doi.org/10.1053/j.semnuclmed.2024.05.005

Dangi, R. R., Sharma, A., & Vageriya, V. (2025). Transforming Healthcare in Low-Resource Settings With Artificial Intelligence: Recent Developments and Outcomes. *Public Health Nurs*, *42*(2), 1017-1030. https://doi.org/10.1111/phn.13500

Durán, J. M., & Jongsma, K. R. (2021). Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *J Med Ethics*. https://doi.org/10.1136/medethics-2020-106820

Elfaham, R. H., Alnaaim, S. A., Maqbul, M. S., *et al.*. (2025). The Human-AI Frontier: An In-Depth Exploration of Physicians' Apprehension and Outlook Towards Artificial Intelligence in Gastroenterology Healthcare. *Gastroenterology & Endoscopy*, *In Press*. https://doi.org/10.1016/j.gande.2025.05.001

Elstein, A. S., & Schwartz, A. (2002). Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. *Bmj*, *324*(7339), 729-732. https://doi.org/10.1136/bmj.324.7339.729

Fehr, J., Citro, B., Malpani, R., *et al.*. (2024). A trustworthy AI reality-check: the lack of transparency of artificial intelligence products in healthcare. *Front Digit Health*, *6*, 1267290. https://doi.org/10.3389/fdgth.2024.1267290

Ferrara, E. (2024). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Sci*, *6*(1), 3. https://doi.org/10.3390/sci6010003

Fogo, A. B., Kronbichler, A., & Bajema, I. M. (2024). AI's Threat to the Medical Profession. *Jama*, *331*(6), 471-472. https://doi.org/10.1001/jama.2024.0018

Fransen, S. J., Kwee, T. C., Rouw, D., *et al.*. (2025). Patient perspectives on the use of artificial intelligence in prostate cancer diagnosis on MRI. *Eur Radiol*, *35*(2), 769-775. https://doi.org/10.1007/s00330-024-11012-y

George, A. S., Baskar, D., & Balaji Srikaanth, P. (2024). The Erosion of Cognitive Skills in the Technological Age: How Reliance on Technology Impacts Critical Thinking, Problem-Solving, and Creativity. *2*, 147-163. https://doi.org/10.5281/zenodo.11671150

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. *Societies*, *15*(1), 6. https://doi.org/10.3390/soc15010006

Gibelli, F., Maio, G., & Ricci, G. (2024). Editorial: Healthcare in the age of sapient machines: physician decision-making autonomy faced with artificial intelligence. Ethical, deontological and compensatory aspects. *Front Med (Lausanne)*, *11*, 1477371. https://doi.org/10.3389/fmed.2024.1477371

Gibson, E. J. (2015). Treating the patient who has the disease. *Tex Heart Inst J*, *42*(3), 193-194. https://doi.org/10.14503/thij-15-5204

Gondocs, D., & Dörfler, V. (2024). AI in medical diagnosis: AI prediction & human judgment. *Artificial Intelligence in Medicine*, *149*, 102769. https://doi.org/10.1016/j.artmed.2024.102769

Guo, W., & Chen, Y. (2025). Investigating Whether AI Will Replace Human Physicians and Understanding the Interplay of the Source of Consultation, Health-Related Stigma, and Explanations of Diagnoses on Patients' Evaluations of Medical Consultations: Randomized Factorial Experiment. *J Med Internet Res*, *27*, e66760. https://doi.org/10.2196/66760

Habli, I., Lawton, T., & Porter, Z. (2020). Artificial intelligence in health care: accountability and safety. *Bull World Health Organ*, *98*(4), 251-256. https://doi.org/10.2471/blt.19.237487

Hannawa, A. F., Wu, A. W., Kolyada, A., *et al.*. (2022). The aspects of healthcare quality that are important to health professionals and patients: A qualitative study. *Patient Education and Counseling*, *105*(6), 1561-1570. https://doi.org/10.1016/j.pec.2021.10.016

Harada, Y., Kawamura, R., Yokose, M., *et al.*. (2024). Definitions and Measurements for Atypical Presentations at Risk for Diagnostic Errors in Internal Medicine: Protocol for a Scoping Review. *JMIR Res Protoc*, *13*, e56933. https://doi.org/10.2196/56933

Harishbhai Tilala, M., Kumar Chenchala, P., Choppadandi, A., *et al.*. (2024). Ethical Considerations in the Use of Artificial Intelligence and Machine Learning in Health Care: A Comprehensive Review. *Cureus*, *16*(6), e62443. https://doi.org/10.7759/cureus.62443

Hassan, M., Kushniruk, A., & Borycki, E. (2024). Barriers to and Facilitators of Artificial Intelligence Adoption in Health Care: Scoping Review. *JMIR Hum Factors*, *11*, e48633. https://doi.org/10.2196/48633

Holmboe, E. S., & Durning, S. J. (2014). Assessing clinical reasoning: moving from in vitro to in vivo. *Diagnosis (Berl)*, *1*(1), 111-117. https://doi.org/10.1515/dx-2013-0029

Hopson, S., Mildon, C., Hassard, K., *et al.*. (2024). Equipping Future Physicians with Artificial Intelligence Competencies through Student Associations. *International Medical Education*, *3*(4), 388-394. https://doi.org/10.3390/ime3040029

Hurley, M. E., Lang, B. H., Kostick-Quenet, K. M., *et al.*. (2025). Patient Consent and The Right to Notice and Explanation of AI Systems Used in Health Care. *Am J Bioeth*, *25*(3), 102-114. https://doi.org/10.1080/15265161.2024.2399828

Janumpally, R., Nanua, S., Ngo, A., *et al.*. (2024). Generative artificial intelligence in graduate medical education. *Front Med (Lausanne)*, *11*, 1525604. https://doi.org/10.3389/fmed.2024.1525604

Jeanty, P. W., Narcisse, M. R., & Crastes Dit Sourd, R. (2024). Editorial: Artificial intelligence, machine learning, and data-mining techniques to increase cost-effectiveness in healthcare. *Front Public Health*, *12*, 1525628. https://doi.org/10.3389/fpubh.2024.1525628

Jha, D., Durak, G., Sharma, V., *et al.*. (2025). A Conceptual Framework for Applying Ethical Principles of AI to Medical Practice. *Bioengineering (Basel)*, *12*(2), 180. https://doi.org/10.3390/bioengineering12020180

Jutel, A. (2009). Sociology of diagnosis: a preliminary review. *Sociol Health Illn*, *31*(2), 278-299. https://doi.org/10.1111/j.1467-9566.2008.01152.x

Kaushal, A., Altman, R., & Langlotz, C. (2020). Geographic Distribution of US Cohorts Used to Train Deep Learning Algorithms. *Jama*, *324*(12), 1212-1213. https://doi.org/10.1001/jama.2020.12067

Khan, M. S., Umer, H., & Faruqe, F. (2024). Artificial intelligence for low income countries. *Humanities and Social Sciences Communications*, *11*(1), 1422. https://doi.org/10.1057/s41599-024-03947-w

Khanna, N. N., Maindarkar, M. A., Viswanathan, V., *et al.*. (2022). Economics of Artificial Intelligence in Healthcare: Diagnosis vs. Treatment. *Healthcare (Basel)*, *10*(12), 2493. https://doi.org/10.3390/healthcare10122493

King, D., Nori, H., & Microsoft. (2025, 30 June). *The Path to Medical Superintelligence*. Microsoft AI. Retrieved 2 July 2025 from, https://microsoft.ai/new/the-path-to-medical-superintelligence/

Knopp, M. I., Warm, E. J., Weber, D., *et al.*. (2023). AI-Enabled Medical Education: Threads of Change, Promising Futures, and Risky Realities Across Four Potential Future Worlds. *JMIR Med Educ*, *9*, e50373. https://doi.org/10.2196/50373

Krishnan, G., Singh, S., Pathania, M., *et al.*. (2023). Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm. *Front Artif Intell*, *6*, 1227091. https://doi.org/10.3389/frai.2023.1227091

Krügel, S., Ammeling, J., Aubreville, M., *et al.*. (2025). Perceived responsibility in AI-supported medicine. *AI & SOCIETY*, *40*(3), 1485-1495. https://doi.org/10.1007/s00146-024-01972-6

Kundu, S. (2021). How will artificial intelligence change medical training? *Communications Medicine*, *1*(1), 8. https://doi.org/10.1038/s43856-021-00003-5

Lenharo, M. (2025). Google AI better than human doctors at diagnosing rashes from pictures. *Nature*. https://doi.org/10.1038/d41586-025-01437-w

Li, Y., & Li, J. (2024). Generative artificial intelligence in medical education: way to solve the problems. *Postgrad Med J*, *100*(1181), 203-204. https://doi.org/10.1093/postmj/qgad116

Limb, M. (2025). Microsoft claims AI tool can outperform doctors in diagnostic accuracy. *Bmj*, *390*, r1385. https://doi.org/10.1136/bmj.r1385

Maleki Varnosfaderani, S., & Forouzanfar, M. (2024). The Role of AI in Hospitals and Clinics: Transforming Healthcare in the 21st Century. *Bioengineering (Basel)*, *11*(4), 337. https://doi.org/10.3390/bioengineering11040337

Mamary, A. J., Stewart, J. I., Kinney, G. L., *et al.*. (2018). Race and Gender Disparities are Evident in COPD Underdiagnoses Across all Severities of Measured Airflow Obstruction. *Chronic Obstr Pulm Dis*, *5*(3), 177-184. https://doi.org/10.15326/jcopdf.5.3.2017.0145

Martínez-Vargas, E., Mora-Jiménez, J., Arguedas-Chacón, S., *et al.*. (2025). The Emerging Role of Artificial Intelligence in Dermatology: A Systematic Review of Its Clinical Applications. *Dermato*, *5*(2), 9. https://doi.org/10.3390/dermato5020009

McDuff, D., Schaekermann, M., Tu, T., *et al.*. (2025). Towards accurate differential diagnosis with large language models. *Nature*, *642*(8067), 451-457. https://doi.org/10.1038/s41586-025-08869-4

Mennella, C., Maniscalco, U., De Pietro, G., *et al.*. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*, *10*(4), e26297. https://doi.org/10.1016/j.heliyon.2024.e26297

Milmo, D. (2025, 30 June). *Microsoft says AI system better than doctors at diagnosing complex health conditions*. Guardian News & Media Limited. Retrieved 2 July 2025, from https://www.theguardian.com/technology/2025/jun/30/microsoft-ai-system-better-doctors-diagnosing-health-conditions-research

Miranda, J., Pereira-Silva, R., Guichard, J., *et al.*. (2025). Artificial Intelligence Outperforms Physicians in General Medical Knowledge, Except in the Paediatrics Domain: A Cross-Sectional Study. *Bioengineering (Basel)*, *12*(6), 653. https://doi.org/10.3390/bioengineering12060653

Mittermaier, M., Raza, M. M., & Kvedar, J. C. (2023). Bias in AI-based models for medical applications: challenges and mitigation strategies. *npj Digital Medicine*, *6*(1), 113. https://doi.org/10.1038/s41746-023-00858-z

Mondal, H. (2025). Ethical engagement with artificial intelligence in medical education. *Adv Physiol Educ*, *49*(1), 163-165. https://doi.org/10.1152/advan.00188.2024

Moreno-Sánchez, P., Del Ser, J., Gils, M., *et al.*. (2025). A Design Framework for operationalizing Trustworthy Artificial Intelligence in Healthcare: Requirements, Tradeoffs and Challenges for its Clinical Adoption. *arXiv*. https://doi.org/10.48550/arXiv.2504.19179

Naik, N., Hameed, B. M. Z., Shetty, D. K., *et al.*. (2022). Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?. *Front Surg*, *9*, 862322. https://doi.org/10.3389/fsurg.2022.862322

Natali, C., Marconi, L., Duran, L., *et al.*. (2025). AI-Induced Deskilling in Medicine: A Mixed Method Literature Review for Setting a New Research Agenda. *SSRN*. https://doi.org/10.2139/ssrn.5166364

Nichol, A. A., Halley, M. C., Federico, C. A., *et al.*. (2023). Not in my AI: Moral engagement and disengagement in health care AI development. *Pac Symp Biocomput*, *28*, 496-506.

Nori, H., Daswani, M., Kelly, C., *et al.*. (2025). Sequential Diagnosis with Language Models. *arXiv*. https://doi.org/10.48550/arXiv.2506.22405

Oliver, N., Calvard, T., & Potočnik, K. (2017). Cognition, Technology, and Organizational Limits: Lessons from the Air France 447 Disaster. *Organization Science*, *28*(4), 729-743. https://doi.org/10.1287/orsc.2017.1138

Pelaccia, T., Tardif, J., Triby, E., *et al*.. (2015). Insights into emergency physicians' minds in the seconds before and into a patient encounter. *Intern Emerg Med*, *10*(7), 865-873. https://doi.org/10.1007/s11739-015-1283-8

Pregowska, A., & Perkins, M. (2024). Artificial intelligence in medical education: Typologies and ethical approaches. *Ethics & Bioethics*, *14*, 96-113. https://doi.org/10.2478/ebce-2024-0004

Preiksaitis, C., & Rose, C. (2023). Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review. *JMIR Med Educ*, *9*, e48785. https://doi.org/10.2196/48785

Quinn, T. P., Senadeera, M., Jacobs, S., *et al.*. (2021). Trust and medical AI: the challenges we face and the expertise needed to overcome them. *J Am Med Inform Assoc*, *28*(4), 890-894. https://doi.org/10.1093/jamia/ocaa268

Rawat, A. S., Rajendran, J., & Sikarwar, S. S. (2025). Introduction to AI in Disease Detection — An Overview of the Use of AI in Detecting Diseases, Including the Benefits and Limitations of the Technology. In *AI in Disease Detection* (pp. 1-26). https://doi.org/10.1002/9781394278695.ch1

Reddy, S. (2024). Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. *Implementation Science*, *19*(1), 27. https://doi.org/10.1186/s13012-024-01357-9

Rodler, S., Kopliku, R., Ulrich, D., *et al.*. (2024). Patients' Trust in Artificial Intelligence-based Decision-making for Localized Prostate Cancer: Results from a Prospective Trial. *Eur Urol Focus*, *10*(4), 654-661. https://doi.org/10.1016/j.euf.2023.10.020

Rony, M. K. K., Parvin, M. R., Wahiduzzaman, M., *et al.*. (2024). "I Wonder if my Years of Training and Expertise Will be Devalued by Machines": Concerns About the Replacement of Medical Professionals by Artificial Intelligence. *SAGE Open Nurs*, *10*, 23779608241245220. https://doi.org/10.1177/23779608241245220

Rosenberg, C. E. (2002). The tyranny of diagnosis: specific entities and individual experience. *Milbank Q*, *80*(2), 237-260. https://doi.org/10.1111/1468-0009.t01-1-00003

Rubin, M., Li, J. Z., Zimmerman, F., *et al.*. (2025). Comparing the value of perceived human versus AI-generated empathy. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-025-02247-w

Safranek, C. W., Sidamon-Eristoff, A. E., Gilson, A., *et al.*. (2023). The Role of Large Language Models in Medical Education: Applications and Implications. *JMIR Med Educ*, *9*, e50945. https://doi.org/10.2196/50945

Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare (Basel)*, *11*(6), 887. https://doi.org/10.3390/healthcare11060887

Sallam, M. (2024). Bibliometric top ten healthcare-related ChatGPT publications in the first ChatGPT anniversary. *Narra J*, *4*(2), e917. https://doi.org/10.52225/narra.v4i2.917

Sallam, M., Al-Mahzoum, K., Alaraji, H., *et al.*. (2025a). Apprehension toward generative artificial intelligence in healthcare: a multinational study among health sciences students [Original Research]. *Frontiers in Education*, *10*, 1542769. https://doi.org/10.3389/feduc.2025.1542769

Sallam, M., Al-Mahzoum, K., Almutairi, Y. M., *et al.*. (2024). Anxiety among Medical Students Regarding Generative Artificial Intelligence Models: A Pilot Descriptive Study. *International Medical Education*, *3*(4), 406-425. https://doi.org/10.3390/ime3040031

Sallam, M., Al-Mahzoum, K., Sallam, M., *et al.*. (2025b). DeepSeek: Is it the End of Generative AI Monopoly or the Mark of the Impending Doomsday?. *Mesopotamian Journal of Big Data*, *2025*, 26-34. https://doi.org/10.58496/MJBD/2025/002

Sallam, M., Al-Salahat, K., Eid, H., *et al.*. (2024). Human versus Artificial Intelligence: ChatGPT-4 Outperforming Bing, Bard, ChatGPT-3.5 and Humans in Clinical Chemistry Multiple-Choice Questions. *Adv Med Educ Pract*, *15*, 857-871. https://doi.org/10.2147/amep.S479801

Sallam, M., Salim, N. A., Barakat, M., *et al.*. (2023). ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J*, *3*(1), e103. https://doi.org/10.52225/narra.v3i1.103

Sarafis, P., Tsounis, A., Malliarou, M., *et al.*. (2013). Disclosing the truth: a dilemma between instilling hope and respecting patient autonomy in everyday clinical practice. *Glob J Health Sci*, *6*(2), 128-137. https://doi.org/10.5539/gjhs.v6n2p128

Schmidt, E., Putora, P. M., & Fijten, R. (2025). The Epistemic Cost of Opacity: How the Use of Artificial Intelligence Undermines the Knowledge of Medical Doctors in High-Stakes Contexts. *Philosophy & Technology*, *38*(1), 5. https://doi.org/10.1007/s13347-024-00834-9

Schubert, T., Oosterlinck, T., Stevens, R. D., *et al.*. (2025). AI education for clinicians. *EClinicalMedicine*, *79*, 102968. https://doi.org/10.1016/j.eclinm.2024.102968

Schuitmaker, L., Drogt, J., Benders, M., *et al.*. (2025). Physicians' required competencies in AI-assisted clinical settings: a systematic review. *Br Med Bull*, *153*(1). https://doi.org/10.1093/bmb/ldae025

Selçuk, S., Kurt Konca, N., & Kaya, S. (2025). AI-driven civil litigation: Navigating the right to a fair trial. *Computer Law & Security Review*, *57*, 106136. https://doi.org/10.1016/j.clsr.2025.106136

Seyyed-Kalantari, L., Zhang, H., McDermott, M. B. A., *et al.*. (2021). Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med*, *27*(12), 2176-2182. https://doi.org/10.1038/s41591-021-01595-0

Shafi, S., & Parwani, A. V. (2023). Artificial intelligence in diagnostic pathology. *Diagnostic Pathology*, *18*(1), 109. https://doi.org/10.1186/s13000-023-01375-z

Shrank, W. H., Rogstad, T. L., & Parekh, N. (2019). Waste in the US Health Care System: Estimated Costs and Potential for Savings. *Jama*, *322*(15), 1501-1509. https://doi.org/10.1001/jama.2019.13978

Shumway, D. O., & Hartman, H. J. (2024). Medical malpractice liability in large language model artificial intelligence: legal review and policy recommendations. *J Osteopath Med*, *124*(7), 287-290. https://doi.org/10.1515/jom-2023-0229

Siala, H., & Wang, Y. (2022). SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. *Social Science & Medicine*, *296*, 114782. https://doi.org/10.1016/j.socscimed.2022.114782

Singh, B., & Kaunert, C. (2025). Accelerating Healthcare Outcomes Uplifting Patient Experiences Pairing Artificial intelligence (AI): Transforming the Healthcare Industry in the Digital Arena. In N. Wickramasinghe (Ed.), *Impact of Digital Solutions for Improved Healthcare Delivery* (pp. 121-146). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-5237-3.ch004

Singh, M. P., & Keche, Y. N. (2025). Ethical Integration of Artificial Intelligence in Healthcare: Narrative Review of Global Challenges and Strategic Solutions. *Cureus*, *17*(5), e84804. https://doi.org/10.7759/cureus.84804

Sun, Y., Sheng, D., Zhou, Z., *et al.*. (2024). AI hallucination: towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanities and Social Sciences Communications*, *11*(1), 1278. https://doi.org/10.1057/s41599-024-03811-x

Taherdoost, H., & Ghofrani, A. (2024). AI's role in revolutionizing personalized medicine by reshaping pharmacogenomics and drug therapy. *Intelligent Pharmacy*, *2*. https://doi.org/10.1016/j.ipha.2024.08.005

Takita, H., Kabata, D., Walston, S. L., *et al.*. (2025). A systematic review and meta-analysis of diagnostic performance comparison between generative AI and physicians. *npj Digital Medicine*, *8*(1), 175. https://doi.org/10.1038/s41746-025-01543-z

Thompson, R. A. M., Shah, Y. B., Aguirre, F., *et al.*. (2025). Artificial Intelligence Use in Medical Education: Best Practices and Future Directions. *Curr Urol Rep*, *26*(1), 45. https://doi.org/10.1007/s11934-025-01277-1

Tigard, D. W. (2021). Artificial Moral Responsibility: How We Can and Cannot Hold Machines Responsible. *Camb Q Healthc Ethics*, *30*(3), 435-447. https://doi.org/10.1017/s0963180120000985

Tranberg, M., Ekedahl, H., Fürst, C. J., *et al.*. (2024). The influence of "bad news" and "neutral/good news" on patients' perception of physician empathy during oncology consultations. *Cancer Med*, *13*(1), e6903. https://doi.org/10.1002/cam4.6903

Tsang, R. (2023). Practical Applications of ChatGPT in Undergraduate Medical Education. *J Med Educ Curric Dev*, *10*, 23821205231178449. https://doi.org/10.1177/23821205231178449

Varkey, B. (2021). Principles of Clinical Ethics and Their Application to Practice. *Med Princ Pract*, *30*(1), 17-28. https://doi.org/10.1159/000509119

Wang, C., Liu, S., Yang, H., *et al.*. (2023). Ethical Considerations of Using ChatGPT in Health Care. *J Med Internet Res*, *25*, e48009. https://doi.org/10.2196/48009

Wei, X., Kumar, N., & Zhang, H. (2025). Addressing bias in generative AI: Challenges and research opportunities in information management. *Information & Management*, *62*(2), 104103. https://doi.org/10.1016/j.im.2025.104103

Xu, X., Chen, Y., & Miao, J. (2024). Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: a systematic scoping review. *J Educ Eval Health Prof*, *21*, 6. https://doi.org/10.3352/jeehp.2024.21.6

Yim, D., Khuntia, J., Parameswaran, V., *et al.*. (2024). Preliminary Evidence of the Use of Generative AI in Health Care Clinical Services: Systematic Narrative Review. *JMIR Med Inform*, *12*, e52073. https://doi.org/10.2196/52073

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, *11*(1), 28. https://doi.org/10.1186/s40561-024-00316-7