Croat Med J. 2025;66:243-6
https://doi.org/10.3325/cmj.2025.66.243

The promises and perils of artificial intelligence in clinical medicine: a dual perspective from internal medicine and pediatrics

Lovro Lamot[1,2], Robert Likić[1,3]

[1]University of Zagreb School of Medicine, Zagreb, Croatia

[2]Department of Pediatrics, Division of Nephrology, Dialysis and Transplantation, University Hospital Centre Zagreb, Zagreb, Croatia

[3]Department of Internal Medicine, Division of Clinical Pharmacology, University Hospital Centre Zagreb, Zagreb, Croatia

robert.likic@mef.hr

The integration of artificial intelligence (AI) into clinical medicine represents one of the most profound paradigm shifts in the history of health care. Its proponents emphasize AI's capacity to process and synthesize vast, heterogeneous clinical data, offering support for clinical decision-making, enhancing pharmacotherapeutic precision, and even accelerating translational research (1-3). At the same time, many clinicians remain appropriately cautious, recognizing AI's current limitations, its potential for error propagation, and the risk of undermining clinical judgment if not properly governed (4,5). While much of the published debate remains speculative or theoretical, real-world clinical practice is increasingly intersecting with AI applications. This article offers a joint perspective from two actively practicing physicians, one in internal medicine and clinical pharmacology and the other in pediatrics, each drawing directly from personal experience with AI-powered clinical support systems such as ChatGPT-4o.

In the field of internal medicine and clinical pharmacology, the potential for AI to transform daily practice is already becoming reality. A clear example is its role in managing complex polypharmacy and suspected adverse drug reactions, where clinical pharmacologists often confront highly intricate medication lists involving both prescription drugs and over-the-counter supplements (6). One case involved a 74-year-old woman presenting with chronic generalized pruritus persisting for over two years. She had multiple comorbidities, including Hashimoto's thyroiditis, non-insulin-dependent diabetes mellitus, hypertension, and chronic constipation. Her medication list exceeded fifteen agents, encompassing both pharmacologic treatments and numerous supplements such as dehydroepiandrosterone, hydrocortisone, omega-3 fatty acids, zinc, magnesium, vitamin B complex, coenzyme Q10, and phytoestrogens. Traditional medication reconciliation in such cases is time-consuming, and the identification of potential delayed hypersensitivity reactions or iatrogenic contributors to her symptoms requires integration of multiple pharmacological domains.

By leveraging ChatGPT-4o, a comprehensive medication assessment was completed within minutes. AI assistance rapidly synthesized pharmacokinetic pathways, cross-reactivity profiles, and rare hypersensitivity reports, while also proposing a rational protocol for systematic withdrawal and challenge to elucidate potential offending agents (7). The system rapidly consulted pharmacovigilance databases and cross-referenced current literature, generating a depth of analysis that would traditionally require several hours or days. Importantly, while AI provided exhaustive and up-to-date suggestions, clinical oversight remained indispensable in prioritizing drug discontinuation order and closely monitoring patient response. Nevertheless, the efficiency and breadth of AI-supported analysis fundamentally elevated the clinician's capacity to navigate such complex pharmacotherapeutic scenarios.

AI also proved transformative in situations of complex differential diagnosis, as illustrated in the manage-

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC-ND 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ment of a critically ill adult woman presenting with multiorgan failure, requiring mechanical ventilation, intermittent hemodialysis, antifungal therapy for candidiasis, and intensive supportive care. The central diagnostic dilemma revolved around distinguishing between multisystem inflammatory syndrome in adults (MIS-A) and systemic lupus erythematosus, as both entities present with overlapping inflammatory, cardiac, renal, hepatic, and hematological involvement (8). Through iterative consultations, the AI-assisted system generated differential diagnostic frameworks, identified gaps in the existing laboratory workup, and proposed organ-specific therapeutic interventions (9,10). In this context, AI served as a cognitive amplifier for the multidisciplinary team, providing access to rapidly evolving literature on MIS-A phenotypes, recent case series, and evidence-based immunosuppressive strategies. Furthermore, AI-assisted pharmacology modules provided drug-drug interaction assessments, suggesting alternative nephroprotective strategies to minimize further renal injury when considering diuretic therapy modifications such as replacing furosemide. Once again, human expertise remained central, but the AI-powered system significantly expanded the diagnostic bandwidth and improved therapeutic precision in real time.

Perhaps one of the most powerful examples of AI's emerging role lies in the application of pharmacogenomics to guide personalized therapy. In a case involving a 16.5-year-old girl suffering from chronic anxiety and nausea, complicated by poor tolerability to multiple prior psychotropic regimens, pharmacogenomic profiling revealed actionable variants in *CYP2C19* and *CYP2D6*. These findings raised critical concerns regarding the metabolism of several first-line selective serotonin reuptake inhibitors (11). The AI system rapidly integrated the guidelines of Clinical Pharmacogenetics Implementation Consortium, Food and Drug Administration labeling, developmental pharmacokinetics, and recent pharmacogenomic literature to generate a ranked list of viable antidepressant options with individualized dosing recommendations (12,13). Furthermore, it allowed dynamic modeling of potential serotonergic side effects based on predicted plasma concentrations. This capacity for an individualized approach in the context of adolescent psychiatric therapy would be extraordinarily difficult to achieve without AI-assisted computational analysis.

Beyond individual patient care, AI is having an increasingly helpful role in medical education, research design, and wound care management. So far, it has been used to create comprehensive multiple-choice question banks for student assessments, to build interactive, case-based teaching scenarios, and to support *in silico* drug repurposing efforts (14). In wound care, AI-driven visual analysis has also contributed by tracking wound healing over time through serial photograph analysis, thereby offering additional insights that can aid clinicians in tailoring treatment protocols according to the wound healing progress (15).

Still, despite these promising developments, experiences in pediatric medicine continue to highlight important reasons for caution. The complexity of pediatric care, where developmental stages and weight-based variations introduce significant variability, poses particular challenges for AI systems (16). In routine pediatric renal ultrasound assessments, for instance, AI has been able to quickly measure parameters such as renal length, cortical thickness, and parenchymal echogenicity. However, several reports have shown that AI sometimes struggles to distinguish between harmless developmental variations and true pathology, occasionally flagging normal findings as potentially abnormal (17). Although these discrepancies may seem minor, they carry the risk of prompting unnecessary tests or interventions if the AI's suggestions are followed without careful human oversight.

A particularly instructive case underscoring the current limitations of AI in pediatrics involved a 13-year-old girl presenting with persistent fatigue, low-grade fever, newly diagnosed hypertension, elevated erythrocyte sedimentation rate and C-reactive protein, a non-specific rash, and an unremarkable immunologic screen including negative antinuclear antibodies and antineutrophil cytoplasmic antibodies, and normal complement levels. Notably, plasma renin activity was elevated while aldosterone remained within reference range, a constellation that initially raised suspicion for secondary hypertension due to renal or renovascular pathology. Subsequent imaging, including Doppler ultrasonography and CT angiography, failed to reveal any structural or vascular abnormalities. When entered into the AI system, the differential diagnosis was dominated by endocrine causes, renal parenchymal disease, and infectious etiologies. Despite the presence of systemic inflammation and unexplained hypertension, vasculitis was not included among the proposed options. Even after expanding the clinical data set with elevated renin, imaging findings, and detailed symptom chronology, AI systems failed to surface vasculitis as a plausible consideration. This omission likely stemmed from a mechanistic overreliance on classification criteria, which typically include immunologic

serologies, specific imaging findings, or biopsy-confirmed vascular inflammation. In reality, the most plausible diagnosis was only clinically entertained after further evolution of symptoms and persistence of diagnostic ambiguity. The AI system's failure to recognize this entity illustrates a deeper limitation: a lack of clinical anticipation. As noted in recent literature, many AI systems still function as uninterpretable "black boxes," which makes it challenging to explain or validate their diagnostic reasoning or omissions (18). Unlike experienced pediatric subspecialists who intuitively weigh evolving symptom trajectories and remain alert to partial or incomplete phenotypes, current AI models are structurally incapable of forming provisional hypotheses in the absence of fully realized diagnostic features. This case exemplifies a broader concern: in rare pediatric immunologically mediated diseases, especially those lacking standardized guidelines or widely validated diagnostic pathways, AI tools may underperform or even mislead if not critically supervised. The failure was not due to lack of data processing but due to a fundamental deficit in interpretive foresight, the type of clinical reasoning that allows for dynamic hypothesis generation, pattern recognition in evolution, and judgment under uncertainty. Until AI systems can incorporate probabilistic reasoning across time and uncertainty, particularly in children where phenotypes are fluid and diagnostic criteria underdeveloped, their role must remain strictly supportive, never substitutive.

Even in relatively straightforward tasks in pediatric nephrology, such as unit conversions and reference value calculations for metabolic urine analysis, AI tools can yield inconsistent outputs. Such calculations are based on established formulas and reference standards that include factors like body height, weight, body surface area, 24-hour urine volume, creatinine excretion, and concentrations of various solutes. While AI systems should simplify these computations and reduce the time needed for patient management, the outputs frequently contain mistakes due to incorrect application of age-specific norms, faulty unit conversions, or failure to properly adjust for body size. Therefore, these issues often require physicians to review and manually correct the results to ensure patient safety and best health outcomes. This exposes a key limitation of the current AI systems: mistakes can occur unless the system is specifically adapted to the detailed requirements of pediatric practice.

Furthermore, AI-generated advice on medication dosing in pediatric patients has so far demonstrated several shortcomings. For instance, algorithms that calculate doses based on body weight have occasionally failed to adjust for organ changes with age, particularly in premature infants, newborns, or children with specific syndromes (19). This becomes even more important when prescribing drugs with a narrow therapeutic index, where even slight miscalculations in dosing can lead to significant clinical risks. Another important issue is that AI systems often communicate their outputs with a tone of authority, regardless of accuracy. Such confident, yet incorrect, suggestions can lead to automation bias, increasing the odds of clinicians trusting the AI's output without critical review (20). This risk becomes even more pronounced in high-pressure clinical situations, where rapid decisions are often needed (21).

Taken together, these experiences emphasize the central point of our joint statement. While AI brings enormous potential to clinical medicine, it cannot yet substitute for the expert judgment of trained physicians. The more prudent path is to build hybrid models where AI supports and strengthens, rather than replaces, clinical expertise. For safe implementation, there will need to be clear audit mechanisms, transparent validation processes, AI systems tailored to specific specialties, and comprehensive training programs to equip physicians with the necessary AI literacy (22).

In the hands of properly trained clinicians, AI can indeed revolutionize clinical practice by enhancing diagnostic accuracy, personalizing therapy, and accelerating research translation. However, clinicians must remain vigilant stewards of patient safety, critically appraising AI recommendations while using their domain expertise to contextualize machine-generated outputs. The future of AI in medicine is neither utopian nor dystopian. It is hybrid, collaborative, and above all, accountable.

References

1 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44-56. Medline:30617339 doi:10.1038/s41591-018-0300-7

2 Obermeyer Z, Emanuel EJ. Predicting the future - big data, machine learning, and clinical medicine. N Engl J Med. 2016;375:1216-9. Medline:27682033 doi:10.1056/NEJMp1606181

3 Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med. 2019;25:24-9. Medline:30617335 doi:10.1038/s41591-018-0316-z

4 Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20:310. Medline:33256715 doi:10.1186/s12911-020-

01332-6

5 London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep. 2019;49:15-21. Medline:30790315 doi:10.1002/hast.973

6 Salas M, Petracek J, Yalamanchili P, Aimer O, Kasthuril D, Dhingra S, et al. The uuse of artificial intelligence in pharmacovigilance: A systematic review of the literature. Pharmaceut Med. 2022;36:295-306. Medline:35904529

7 Gholap AD, Uddin MJ, Faiyazuddin M, Omri A, Gowri S, Khalid M. Advances in artificial intelligence for drug delivery and development: A comprehensive review. Comput Biol Med. 2024;178:108702. Medline:38878397 doi:10.1016/j.compbiomed.2024.108702

8 Morris SB, Schwartz NG, Patel P, Abbo L, Beauchamps L, Balan S, et al. Case sseries of multisystem inflammatory syndrome in adults associated with SARS-coV-2 infection - United Kingdom and United States, March-August 2020. MMWR Morb Mortal Wkly Rep. 2020;69:1450-6. Medline:33031361 doi:10.15585/mmwr.mm6940e1

9 Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature. 2020;580:252-6. Medline:32269341 doi:10.1038/s41586-020-2145-8

10 Hauben M, Hartford CG. Artificial iintelligence in pharmacovigilance: scoping points to consider. Clin Ther. 2021;43:372-9. Medline:33478803 doi:10.1016/j.clinthera.2020.12.014

11 Hicks JK, Bishop JR, Sangkuhl K, Müller DJ, Ji Y, Leckband SG, et al. Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline or CYP2D6 and CYP2C19 genotypes and dosing of selective serotonin reuptake inhibitors. Clin Pharmacol Ther. 2015;98:127-34. doi:10.1002/cpt.147. Published online Jun 29, 2015. Medline:25974703

12 Luzum JA, Pakyz RE, Elsey AR, Haidar CE, Peterson JF, Whirl-Carrillo M, et al. The pharmacogenomics research network translational pharmacogenetics program: outcomes and metrics of pharmacogenetic implementations across diverse healthcare systems. Clin Pharmacol Ther. 2017;102:502-10. Medline:28090649 doi:10.1002/cpt.630

13 Roden DM, McLeod HL, Relling MV, Williams MS, Mensah GA, Peterson JF, et al. Pharmacogenomics. Lancet. 2019;394:521-32. Medline:31395440 doi:10.1016/S0140-6736(19)31276-0

14 Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. Acad Med. 2018;93:1107-9. Medline:29095704 doi:10.1097/ACM.0000000000002044

15 Misir A. Artificial intelligence and machine learning in diabetic foot ulcer care: advances in diagnosis, treatment, prognosis, and novel therapeutic strategies. J Diabetes Sci Technol. 2025. Medline:40754782 doi:10.1177/19322968251363632

16 Rosen JM. How artificial intelligence, virtual reality, and other digital technologies are changing the field of pediatric neurogastroenterology. Gastroenterol Clin North Am. 2025;54:667-79. Medline:40752924 doi:10.1016/j.gtc.2025.01.001

17 Nagar A, Gobburu J, Chakravarty A. Artificial intelligence in pharmacovigilance: advancing drug safety monitoring and regulatory integration. Ther Adv Drug Saf. 2025;16:20420986251361435. Medline:40756610 doi:10.1177/20420986251361435

18 Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med. 2022;28:31-8. Medline:35058619 doi:10.1038/s41591-021-01614-0

19 Zhang Y, Wei X, Bajaj G, Barrett JS, Meibohm B, Joshi A, et al. Challenges and considerations for development of therapeutic proteins in pediatric patients. J Clin Pharmacol. 2015;55 Suppl 3:S103-15. Medline:25707958 doi:10.1002/jcph.382

20 Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics. 2020;46:205-11. Medline:31748206 doi:10.1136/medethics-2019-105586

21 Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc. 2012;19:121-7. Medline:21685142 doi:10.1136/amiajnl-2011-000089

22 Danysz K, Cicirello S, Mingle E, Assuncao B, Tetarenko N, Mockute R, et al. Artificial intelligence and the future of the drug safety professional. Drug Saf. 2019;42:491-7. Medline:30343417 doi:10.1007/s40264-018-0746-z