EDITORIAL

DOI: http://dx.doi.org/10.15403/jgld-6709

Beyond the Hype: Understanding the Limits, Errors and Risk Areas of Artificial Intelligence in Gastroenterology

Ștefan Lucian Popa

2nd Internal Medicine Department, "Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania

Address for correspondence:
Ștefan Lucian Popa
2nd Internal Medicine Department, "Iuliu Hatieganu" University of Medicine and Pharmacy, Cluj-Napoca, Romania
popa.stefan@umfcluj.ro

Received: 25.11.2025
Accepted: 27.11.2025

Artificial intelligence (AI) has undergone substantial and accelerating development in gastroenterology over the past decade, driven by concurrent advances in computational hardware and algorithmic methodologies, including machine learning, deep learning, and computer vision. Its applications now span nearly the entire clinical spectrum, from real-time endoscopic polyp detection to prediction of inflammatory bowel disease activity (IBD), fibrosis staging in hepatology, automatic capsule endoscopy lesion identification, radiologic and radiomics-based liver and pancreatic disease analysis, digital histopathology for dysplasia and inflammation grading, multimodal prediction of treatment response and disease complications, electronic health record-based clinical decision support systems that assist in risk stratification, workflow optimization, and precision-medicine applications [1-4].

Several computer-aided detection systems have already achieved regulatory approval, and many endoscopy platforms now incorporate AI-based assistance as part of their core functionality. This rapid evolution has generated optimism, with AI frequently portrayed as the next essential tool to the practicing gastroenterologist. Yet, despite the impressive capabilities of modern algorithms, their integration into clinical practice is accompanied by limitations, blind spots, and risks that are often underappreciated. Recognizing where AI systems fail, why they fail, and how those failures may affect patient care is essential to ensure that the technology contributes meaningfully and safely to clinical practice.

Much of the success of AI rests on the quality of the datasets used for training and validation. Most available models in gastroenterology depend on large volumes of endoscopic images, histopathology slides, or electronic health record (EHR) data [1]. However, the representativeness of these datasets is frequently limited. A substantial portion originates from high-volume, technologically advanced academic centers, where imaging quality, bowel preparation, and operator expertise may not reflect real-world conditions [1-3]. Artificial intelligence systems trained in such environments may perform poorly in community hospitals that use older endoscopy platforms, have lower-resolution imaging, or treat patients without the pharmacologic therapies recommended by the international guidelines [5-9]. Sampling and selection biases thereby restrict the generalizability of these tools. Annotation bias adds another layer of complexity [10, 11]. Even when expert specialists provide ground-truth labels, interobserver variability persists for many tasks, including dysplasia grading, inflammatory activity scoring, and characterization of subtle lesions [5-10]. Artificial intelligence models inherit these inconsistencies, and the resulting performance variations are rarely acknowledged in promotional narratives [10-13].

The underrepresentation of rare or atypical findings introduces further weaknesses. Most endoscopic AI systems are trained predominantly on common lesions such as typical adenomas, while rare abnormalities, such as, early gastric cancers, granular cell tumors, ischemic patterns, appear only sporadically in training sets [14-17]. Because deep learning models excel at recognizing patterns they have frequently encountered, they struggle with uncommon presentations, making false negatives more likely in precisely those scenarios where expert assistance would be most valuable [14-17].

Endoscopy remains the clinical area in which AI tools are most developed. Computer-aided detection and diagnosis systems for colorectal lesions demonstrate high sensitivity in controlled studies, yet their limitations become apparent during routine procedures. False positives constitute one of the most frequent operational challenges. Mucosal folds, vascular patterns, bubbles, residual stools, and light reflections often trigger alarms, which can distract the endoscopist and potentially prolong procedures [14-17]. Excessive alerts encourage "alert fatigue," reducing the clinician's ability to distinguish true findings from noise. Conversely, false negatives remain a critical problem, particularly for flat or depressed lesions that are inherently difficult to detect [14-17]. Poor bowel preparation, rapid scope movement, partial occlusion of lesions, and suboptimal lighting further exacerbate this issue. Although sensitivity metrics reported in clinical trials are promising, real-time performance in diverse settings is considerably more variable.

Continuous exposure to AI-assisted polyp detection appears to influence endoscopist performance when subsequently conducting a standard, non-AI colonoscopy [18]. In this retrospective, multi-center analysis of 1,443 diagnostic colonoscopies performed within the ACCEPT trial framework, adenoma detection rates (ADR) in procedures without AI declined significantly following a 3-month period of regular AI use, decreasing from 28.4% to 22.4% [18]. Multivariable modelling identified prior AI exposure as an independent predictor of reduced ADR, alongside established patient-related factors such as older age and male gender [18]. These findings suggest that routine reliance on AI support may induce behavioral or attentional changes that impair baseline detection performance, highlighting the need to better understand how AI systems interact with operator skill, vigilance, and long-term clinical practice.

Another important limitation is the lack of explainability. Most endoscopic AI systems operate as opaque black boxes, offering little or no insight into why a region of mucosa was flagged [5-10]. Without interpretable reasoning, clinicians cannot evaluate when AI output may be incorrect, nor can they learn from its successes and failures [9]. As a result, the clinician's capacity to exercise informed oversight is hampered, and inappropriate trust or mistrust may develop.

Artificial intelligence systems also typically analyze imaging data in isolation, neglecting contextual clinical information, such as family history, previous colonoscopy results, comorbidities, laboratory parameters, or patient age, that influence the significance of endoscopic findings [7]. An algorithm that identifies polyps without integrating clinical relevance may inadvertently encourage overdiagnosis or overtreatment [7]. Truly holistic clinical reasoning remains a limitation of current AI tools, which are optimized for narrow perception tasks rather than broader diagnostic synthesis.

Beyond endoscopy, predictive and prognostic AI models are increasingly used to assess disease activity, predict flares in inflammatory bowel disease, forecast response to biologic therapy, and estimate the risk of cirrhosis decompensation [2]. Yet these models face significant challenges. Electronic health care data, while extensive, is noisy and incomplete. Coding variability, inconsistent documentation, missing laboratory values, and temporal misalignment introduce substantial noise that can distort model predictions [9]. Moreover, clinical environments evolve with new diagnostic criteria, therapeutic options, and management algorithms appear continuously. Models trained on old data may not accommodate these shifts, leading to performance decay or "model drift" unless systems undergo continuous retraining and monitoring [9]. These limitations translate into distinct risk domains that can affect patient safety. Automation bias represents a major hazard [9]. Clinicians, particularly less experienced ones, may over-rely on AI output and discount their own judgment. If an AI system fails to identify a lesion, the endoscopist may also overlook it, assuming the system would have flagged anything significant. Conversely, an incorrect AI-based prediction about disease severity or risk stratification may influence clinical decisions inappropriately [9]. Training clinicians to remain critically engaged and to view AI as an adjunct - not an authority - is therefore essential.

Workflow integration is another overlooked challenge. Artificial intelligence generated alerts, overlays, and visual cues can interrupt the natural rhythm of an endoscopy, increase cognitive load, or obscure visual information [19-22]. Poorly designed interfaces may paradoxically diminish situational awareness [20-22]. Only through careful human-factors engineering can AI become a seamless component of the procedural environment rather than a source of distraction [18-22].

Ethical, legal, and regulatory considerations further complicate the deployment of AI in gastroenterology. Questions about liability remain unsettled: if an AI tool contributes to a missed cancer, responsibility may be ambiguous [23-29]. Transparency in data provenance, informed consent for inclusion in training datasets, and safeguarding patient privacy across institutions are essential concerns [23-29]. Regulatory pathways, although improving, have not fully adapted to the rapid iterative nature of AI development, particularly for models that require ongoing updates to maintain performance [23-29].

Generalizability remains one of the most persistent challenges across all AI applications [5-10]. Models that perform well in the environment where they were trained may deteriorate significantly when transferred elsewhere due to differences in endoscopy hardware, imaging settings, disease prevalence, and even cultural or dietary patterns that influence gastrointestinal health. External validation, though increasingly emphasized, is still insufficient in many published studies. True real-world robustness, across multiple centers, devices, and populations, remains a high bar that few models have convincingly met.

Explainability is also a central determinant of trust. Clinicians need to understand, not necessarily in exhaustive technical detail, but in practical clinical terms, why a model reached a particular conclusion. Without such insight, responsible decision-making becomes difficult. A future in which AI plays a meaningful role in gastroenterology will require not only more accurate algorithms but also more transparent and interpretable ones.

Addressing these challenges requires a multipronged strategy that emphasizes human–machine synergy. Continuous monitoring of AI performance after deployment is essential, analogous to post-marketing surveillance for pharmaceuticals [5-10]. Systems must be audited regularly for accuracy, false positive rates, false negative patterns, and calibration across different patient cohorts. Mechanisms to detect and correct model drift should be integrated into the life cycle of any AI tool. Multimodal AI systems that integrate endoscopic imaging with clinical, genetic, or biochemical data may offer more comprehensive support for decision-making, though these approaches also introduce additional complexity and concerns regarding interpretability [1-3].

Clinician education is equally important. As AI becomes deeply embedded in endoscopy suites and clinical workflows, gastroenterologists must understand the strengths, weaknesses, and failure modes of the systems they use [3]. Training in AI literacy, such as training in clinical epidemiology or biostatistics, will be essential for future generations of practitioners. Ultimately, the goal is to cultivate a relationship in which clinicians remain firmly in control, using AI as an aid but retaining responsibility for final decisions [3].

The future of AI in gastroenterology holds tremendous promise but realizing that promise requires a shift in perspective. Accuracy metrics alone is insufficient. What matters more is clinical relevance, safety, and reliability in real-world settings. Artificial intelligence systems must demonstrate not only technical excellence in controlled datasets but also resilience across diverse environments, transparency in decision pathways, and tangible improvements in patient outcomes. Understanding the limitations of AI is not an impediment to progress; it is a prerequisite for responsible innovation. By moving beyond the hype and critically examining both the strengths and the vulnerabilities of current systems, the field of gastroenterology can ensure that AI becomes a trustworthy collaborator in delivering high-quality, patient-centered care.

Conflicts of interest: None to declare.

REFERENCES

1. Wiest IC, Bhat M, Clusmann J, et al. Large language models for clinical decision support in gastroenterology and hepatology. Nat Rev Gastroenterol Hepatol. 2025;22(11):773787. doi:10.1038/s41575-025-01108-1
2. Bharadwaj HR, Dahiya DS, Dalal P, et al. Artificial Intelligence in Population-Level Gastroenterology and Hepatology: A Comprehensive Review of Public Health Applications and Quantitative Impact. Dig Dis Sci. 2025. doi:10.1007/s10620-025-09452-7
3. Wu YM, Tang FY, Qi ZX. Multimodal artificial intelligence technology in the precision diagnosis and treatment of gastroenterology and hepatology: Innovative applications and challenges. World J Gastroenterol. 2025;31(38):109802. doi:10.3748/wjg.v31.i38.109802
4. Clusmann J, Balaguer-Montero M, Bassegoda O, et al. The barriers to uptake of artificial intelligence in hepatology and how to overcome them. J Hepatol. 2025;83(6):14101426. doi:10.1016/j.jhep.2025.07.003
5. Ho CK, Asrani SK. Ethics, Bias, and Governance in Artificial Intelligence for Hepatology: Toward Building a Safe and Fair Future. J Clin Exp Hepatol. 2025;15(6):102628. doi:10.1016/j.jceh.2025.102628
6. Hann A, Troya J, Fitting D. Status and limitations of artificial intelligence in colonoscopy. United European Gastroenterol J. 2021;9(5):527533. doi:10.1002/ueg2.12108
7. Ruffle JK, Farmer AD, Aziz Q. Artificial Intelligence-Assisted Gastroenterology: Promises and Pitfalls. Am J Gastroenterol. 2019;114(3):422428. doi:10.1038/s41395-018-0268-4
8. Sutton RA, Sharma P. Overcoming barriers to implementation of artificial intelligence in gastroenterology. Best Pract Res Clin Gastroenterol. 2021;5253:101732. doi:10.1016/j.bpg.2021.101732
9. Koçak B, Ponsiglione A, Stanzione A, et al. Bias in artificial intelligence for medical imaging: fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects. Diagn Interv Radiol. 2024;31(2):7588. doi:10.4274/dir.2024.242854
10. Cao JS, Lu ZY, Chen MY, et al. Artificial intelligence in gastroenterology and hepatology: Status and Challenges. World J Gastroenterol. 2021;27(16):16641690. doi:10.3748/wjg.v27.i16.1664
11. Ibrahim SA, Pronovost PJ. Diagnostic Errors, Health Disparities, and Artificial Intelligence: A Combination for Health or Harm? JAMA Health Forum. 2021;2(9):e212430. doi:10.1001/jamahealthforum.2021.2430
12. Nakashima H, Kawahira H, Kawachi H, et al. Artificial intelligence diagnosis of Helicobacter pylori infection using blue laser imaging-bright and linked color imaging: a single-center prospective study. Ann Gastroenterol. 2018;31(4):462468. doi:10.20524/aog.2018.0269
13. Rao BH, Trieu JA, Nair P, et al. Artificial intelligence in endoscopy: More than what meets the eye in screening colonoscopy and endosonographic evaluation of pancreatic lesions. Artif Intell Gastrointest Endosc. 2022;3(3):1630. doi:10.37126/aige.v3.i3.16
14. von Renteln D. Artificial intelligence for colorectal polyp sizing: clinical promise and methodological challenges. Endoscopy. 2025. doi:10.1055/a-2706-0201
15. Areia M, Mori Y, Correale L, et al. Cost-effectiveness of artificial intelligence for screening colonoscopy: a modelling study. Lancet Digit Health. 2022;4(6):e436e444. doi:10.1016/S2589-7500(22)00042-5
16. Sibomana O, Saka SA, Grace Uwizeyimana M, et al. Artificial Intelligence-Assisted Endoscopy in Diagnosis of Gastrointestinal Tumors: A Review of Systematic Reviews and Meta-Analyses. Gastro Hep Adv. 2025;4(9):100754. doi:10.1016/j.gastha.2025.100754
17. Ramzan M, Raza M, Khan ZF, et al. A review on computer-aided diagnostic system to classify the disorders of the gastrointestinal tract. Eur J Med Res. 2025;30(1):674. doi:10.1186/s40001-025-02718-w
18. Budzyń K, Romańczyk M, Kitala D, et al. Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. Lancet Gastroenterol Hepatol. 2025;10(10):896903. doi:10.1016/S2468-1253(25)00133-5
19. Zhou LQ, Wang JY, Yu SY, et al. Artificial intelligence in medical imaging of the liver. World J Gastroenterol. 2019;25(6):672682. doi:10.3748/wjg.v25.i6.672
20. Rosa B, Mascarenhas MJ, Afonso J, et al. Artificial intelligence-assisted versus conventional reading in pan-intestinal capsule endoscopy for suspected mid-lower gastrointestinal bleeding. BMJ Open Gastroenterol. 2025;12(1):e001906. doi:10.1136/bmjgast-2025-001906
21. Soffer S, Klang E, Shimon O, et al. Deep learning for wireless capsule endoscopy: a systematic review and meta-analysis. Gastrointest Endosc. 2020;92(4):831839.e8. doi:10.1016/j.gie.2020.04.039
22. Ibrahim AF, Danpanichkul P, Hayek A, et al. Artificial Intelligence in Gastroenterology Education: DeepSeek Passes the Gastroenterology Board Examination and Outperforms Legacy ChatGPT Models. Am J Gastroenterol. 2025. doi:10.14309/ajg.0000000000003552
23. Ramoni D, Scuricini A, Carbone F, et al. Artificial intelligence in gastroenterology: Ethical and diagnostic challenges in clinical practice. World J Gastroenterol. 2025;31(10):102725. doi:10.3748/wjg.v31.i10.102725
24. Campion JR, O'Connor DB, Lahiff C. Human-artificial intelligence interaction in gastrointestinal endoscopy. World J Gastrointest Endosc. 2024;16(3):126135. doi:10.4253/wjge.v16.i3.126
25. Montalbano P. Ethical and Legal Considerations of Medical Artificial Intelligence. Prim Care. 2025;52(4):769779. doi:10.1016/j.pop.2025.07.009

26. El-Sayed A, Lovat LB, Ahmad OF. Clinical Implementation of Artificial Intelligence in Gastroenterology: Current Landscape, Regulatory Challenges, and Ethical Issues. Gastroenterology. 2025;169(3):518530. doi:10.1053/j.gastro.2025.01.254
27. Astărăstoae V, Rogozea LM, Leașu F, et al. Ethical Dilemmas of Using Artificial Intelligence in Medicine. Am J Ther. 2024;31(4):e388e397. doi:10.1097/MJT.0000000000001693
28. Aggarwal N, Drew DA, Parikh RB, et al. Ethical Implications of Artificial Intelligence in Gastroenterology: The Co-pilot or the Captain? Dig Dis Sci. 2024;69(8):27272733. doi:10.1007/s10620-024-08557-9
29. Ning Y, Teixayavong S, Shang Y, et al. Generative artificial intelligence and ethical considerations in health care: a scoping review and ethics checklist. Lancet Digit Health. 2024;6(11):e848e856. doi:10.1016/S2589-7500(24)00143-2