



Priorities for advancing artificial intelligence in inflammatory bowel disease: a technology update

Anna L. Silverman, MD,^{1,2,*} Ryan W. Stidham, MS, MD,^{3,4,5,*} Sravanthi Parasa, MD,^{6,*}
Prateek Sharma, MD,^{7,8,*} Uma Mahadevan, MD^{9,*}

New York, New York; Scottsdale, Arizona; Ann Arbor, Michigan; Seattle, Washington; Kansas City, Missouri; San Francisco, California, USA

Artificial intelligence (AI), encompassing various methods that enable computers to process complex data and perform tasks previously requiring human input, is set to play a pivotal role in the future of inflammatory bowel diseases (IBDs) health care and research. This review examines AI's role in IBD and outlines a roadmap of priorities for developments in the field. To fully harness AI's potential for advancing high-quality health care and scientific discovery in IBD, large and diverse data sets are essential, along with AI tools that undergo rigorous and continuous validation. In addition, because AI becomes increasingly integrated into IBD care, it will be essential to study and optimize clinician–AI interactions and to provide education for health care professionals on the effective use of AI. (Gastrointest Endosc 2026;103:661-71.)

INTRODUCTION

Caring for patients with inflammatory bowel disease (IBD) is becoming increasingly complex, particularly with the introduction of multiple new targeted immune modulator (TIM) therapies over the past decade. Despite the expanding range of treatment options, a significant need for improvement in clinical care remains because only a minority of patients achieve the treatment target of mucosal healing.^{1,2} Furthermore, despite efforts to standardize care to a minimum quality level, IBD management is highly variable, with evidence that care is fragmented for many patients across centers and providers.³ In addition, gaps persist in research, particularly in understanding the pathogenesis of IBD, opportunities for early interception of the disease, and determining the optimal allocation of medical and surgical therapies. Because IBD care becomes more specialized, tools that enhance the standardization of specialty care stand to improve population-level outcomes.

Artificial intelligence (AI), which includes multiple methods that enable computers to process complex data and complete tasks that previously required humans, is positioned to potentially be a powerful tool to address these clinical and research gaps. Although AI applications in IBD remain in their early investigational stages, there is emerging evidence that suggests AI will have a widespread impact on both clinical care and scientific advancements.⁴⁻¹⁰ In addition, AI stands to equip clinicians and researchers with new tools and methods for exploring insights and making new

discoveries. This review explores the role of AI in IBD and prioritizes a roadmap for its development in the field (Table 1). Importantly, realizing the full potential of AI to advance high-quality health care delivery and scientific discovery in IBD will require large, diverse, de-identified IBD data sets, along with AI tools that are rigorously and continuously validated. Studying and optimizing clinician–AI interactions will also provide essential guidance because AI becomes embedded in IBD care and research.

INTRODUCTION TO AI

AI incorporates components of electronic data collection, modeling of digital information to represent complex, often qualitative human concepts (eg, “remission”) and ultimately, replicating human cognitive tasks.^{5,11} Table 2 outlines the key AI terms including machine learning (ML), deep learning (DL), natural language processing (NLP), large language models (LLMs), computer vision (CV), and generative AI along with their definitions, hierarchical relationships as subsets within AI, and common applications such as classification, prediction, dynamic language generation, and synthetic data creation.¹²

Regardless of the type of AI used, all models require training data where the AI learns patterns. A subset of data not seen during training is used to test the model after training concludes. A variety of metrics, such as F1 score, which balances the models' ability to correctly identify positive instances and minimize false positives and

TABLE 1. AI in IBD priorities

AI in IBD priority	Unmet needs and priority addresses
Widely adopted automated disease activity monitoring enabled by AI	• Standardized disease activity assessment • Population health • Real-world comparative effectiveness
Decreasing the time from therapeutic discovery to regulatory approval with AI integration	• Decrease trial costs • Decreased time to enrollment • Decrease placebo needs • Increase therapeutics under investigation
Optimize therapeutic selection with AI-driven precision medicine tools	• Decrease uncontrolled disease • Improve clinical outcomes
Improve IBD administrative workflows with AI	• Reduce insurance approval delays • Reduce human effort in administrative tasks • Patient education
Large public de-identified multimodal IBD data sets to maximize the potential of AI in IBD	• Support IBD AI innovation • Allow for benchmarking and comparative performance of models
Regulation to support integration of AI into IBD clinical practice	• Improve algorithm and safety • Advance implementation science of clinician–AI interaction

The page numbers in parentheses reflect the location of each priority discussed in this article.
AI, Artificial intelligence; IBD, inflammatory bowel disease.

negatives, are used to evaluate the performance of models depending on the task.¹³ The field of AI is undergoing rapid advancement with new model architectures that require progressively less training data and hardware usage, increasing the availability of these tools.¹⁴ Importantly, AI is unable to learn from data that do not exist; if a clinician did not capture important findings in their procedure or clinical encounter, AI will not be able to learn from or detect them. There remains a role for note templates to help ensure all important information is captured.¹⁵ This technology update was developed through an expert review process led by members of the American Society for Gastrointestinal Endoscopy AI in Gastroenterology Institute with recognized expertise in AI and IBD. Existing AI applications in IBD, with a view toward the steps needed to bring AI tools from pilot-studies to large-scale evaluation and implementation, are focused. Expert insights were gathered through iterative discussions and a critical review of emerging evidence to identify key challenges, opportunities, and priorities for advancing AI applications in IBD. The final synthesis reflects the expert panel’s collective perspectives, supported by a targeted literature review and consensus refinement.

AI in IBD priority: widely adopted automated disease activity monitoring enabled by AI

Disease activity monitoring through endoscopy, histology, and imaging is critical in the modern treatment to target paradigm of aiming for mucosal normalization.¹⁶ However, the scoring systems used in endoscopy, pathology, and imaging suffer from suboptimal adoption in IBD clinical practice. In addition to limited uptake in routine care, there can be subjectivity inherent in these scores.¹⁷

Automated disease activity assessment is likely the most near-term task in IBD care and research that AI will soon standardize and automate.

Passive AI-driven disease activity scoring in endoscopy, imaging, and pathology offers significant benefits beyond standardization. It enables quicker identification of patients for clinical trials without waiting for central reader confirmation or repeat endoscopy, expediting enrollment.¹⁸ Clinical trials can reduce reliance on resource-intensive human readers, using algorithmic evaluation for most cases and reserving human review for low-confidence scores. Over time, this approach may lower trial costs, expand investigational therapy pipelines, and enhance standardized, reproducible disease assessment in clinical and regulated settings.

Endoscopy and histology

Over the last 2 decades, research groups have successfully shown that AI is capable of performing disease activity assessment using existing standardized scoring systems of endoscopic images, video, and procedure report text with similar accuracy as expert humans, especially in ulcerative colitis (UC) with the Mayo Endoscopic Score (MES) and Ulcerative Colitis Endoscopic Index of Severity.^{11,19-28} A recent systematic review and meta-analysis of 12 studies demonstrated that AI achieved high sensitivity and specificity in assessing mucosal healing in UC, with a sensitivity of 0.91 and a specificity of 0.89 for still images and a sensitivity of 0.86 and a specificity of 0.91 for videos.²⁸ However, the evidence is limited by moderate-to-high heterogeneity between studies (I^2 over 90%). The creation of large data sets is essential to refine AI-driven disease activity assessment, ensuring its

TABLE 2. AI terms, definitions, hierarchical relationships, and applications

Term	Definition	Subset	Applications
Machine learning (ML)	Allows computers to learn from data without being explicitly programmed.	AI	Classification, prediction
Deep learning	Uses algorithm architectures modeled after the human brain to process large data sets.	ML	Classification, prediction
Natural language processing (NLP)	Focuses on extracting information from written and spoken language.	AI	Classification, prediction
Large language models (LLMs)	Aims to analyze conversational patterns and concepts using natural language.	NLP	Generative AI tasks, such as chatbots and dynamic responses
Computer vision	Focuses on processing image data, such as radiology or pathology images.	AI	Classification, prediction
Generative AI	Models that create new content (eg, text, images) inspired by similar data.	AI (includes LLMs)	Synthetic data generation (text, patient records, molecular structures, medical images), dynamic language models

AI, Artificial intelligence.

reliability, standardization, and integration into routine clinical practice and research.

Histologic remission has been proposed as an aspirational goal in the treatment of IBD; however, we are just beginning to understand the relationship between histologic features and disease outcomes.¹⁶ AI analysis of whole slide images is able to identify and quantify disease activity using the Geboes score, a simplified Paddington International virtual ChromoendoScopy ScOre Histologic Remission Index, and the Global Histologic Disease Activity Score, with similar agreement between expert pathologists.²⁹⁻³⁵ For example, Rymarczyk and colleagues³¹ developed AI models for assessing histologic disease activity based on the Global Histology Activity Score (GHAS) for Crohn's disease (CD) and Geboes subgrades for UC. These models achieved accuracies ranging from 80% to 89% for GHAS in the colon, 65% to 82% for GHAS in the ileum, and 65% to 85% for Geboes scale. AI has proven useful in real-time analysis of endocytoscopy, which allows for microscopic evaluation of the mucosa at the time of endoscopy.³⁶⁻³⁹ In addition, AI analysis of histologic activity has shown promise in its ability to predict future flare within a year, which may aid in therapeutic decision-making.^{33,40} These advancements highlight AI's potential to enhance histologic assessment, aiding in a more precise and predictive approach to IBD management.

Imaging

Imaging is an important complementary method for assessing disease activity alongside endoscopy. Currently developed AI tools can determine disease activity scores used in cross-sectional imaging and video capsule endoscopy (VCE).⁴¹⁻⁴⁷ Imaging allows for transmural evaluation of disease activity, which can be seen via CT, MRI, and intestinal US (IUS). Recent advances in CV and DL, particularly neural network and transformer architectures, have shown the ability to quantify structural bowel damage with similar accuracy to expert radiologists but in less

time.⁴⁸ For example, Holland et al⁴⁹ developed an AI disease severity analysis of magnetic resonance enterography images, which detected 100% of severe cases and showed a strong correlation with the existing magnetic resonance index of activity score. IUS, the only imaging modality able to evaluate the bowel wall layers, is leading to new understandings of IBD disease activity as well as serving as a noninvasive point-of-care disease activity assessment for clinical and trial use.⁵⁰⁻⁵² Pilot studies have shown feasibility of AI to automate disease activity assessment for IUS with an area under the curve (AUC) of 0.75 to 0.98.^{53,54} In addition, AI can aid in helping inexperienced providers acquire diagnostic-quality images comparable to experts without AI, as seen in lung US.⁵⁵

AI-based VCE models can replicate expert scoring of CD ulcers with an AUC of 0.91 for the Lewis score ($r = 0.751$; $P < .0001$) and the Capsule Endoscopy Crohn's Disease Activity Index ($r = 0.707$; $P < .0001$).^{44,56} In addition, AI models achieve 91% to 99% accuracy in classifying the MES severity of VCE used to evaluate colitis.^{47,57} Reducing the length of the VCE video via AI selecting only the important frames from a mean of 303 minutes to just over 3 minutes may translate to increased use of VCE in clinical practice and trials.⁵⁸

Current methods of disease evaluation, including imaging and endoscopy, are suboptimal for objectively delineating the ratio of inflammatory versus fibrotic tissue, which has been limiting the development of antifibrotic therapies. Radiomics, AI image analysis of texture, shape, intensity, and spatial relationships within the image that may not be visible to the human eye, may be the tool that allows for objective fibrosis evaluation, supporting the evaluation of future therapeutics. An AI model trained on enterography images and matched stricture surgical resection histology was developed, demonstrating superior accuracy in predicting moderate-to-severe fibrosis than experienced radiologists, with an AUC of 0.75 versus 0.59.⁵⁹ Other groups have confirmed AI's superiority in distinguishing fibrosis over expert radiologists.^{60,61}

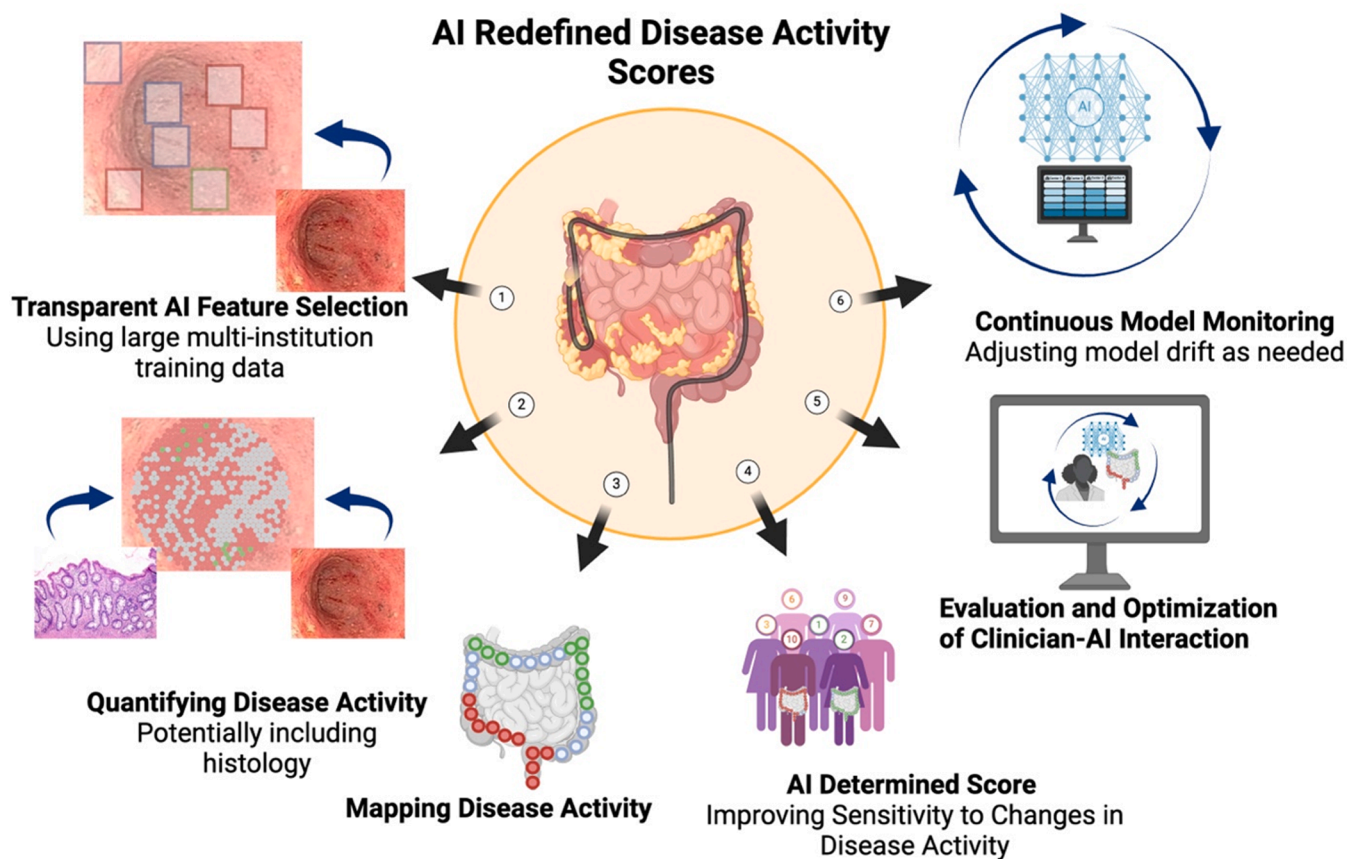


Figure 1. Proposed pathway for development and maintenance of artificial intelligence (AI) redefined disease activity scores.

New AI-enabled measures and future multimodel models of IBD

AI will help the field redefine objective disease activity evaluations by creating new disease activity instruments that are able to capture more granular detail (Fig. 1). A recent example is the work by Stidham et al,⁵ which introduced a novel disease activity score derived from a CV model, known as the cumulative disease score (CDS). CDS quantifies disease activity by summing MES at regular intervals along the bowel, enhancing the ability to differentiate treatment responses among patients. CDS successfully stratified the expected future endoscopic response of patients in the phase 3 Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis trial of Ustekinumab. Moreover, CDS demonstrated greater sensitivity in detecting endoscopic changes than MES (Hedges *g*, 0.743 vs 0.460), indicating that 50% fewer patients are required to detect a therapeutic effect. Another example is the cumulative ileal damage severity score, which is an AI model trained to evaluate disease severity at 1-cm intervals of enterography images, which was superior to conventional methods of imaging disease activity assessment for predicting TIM failure in a pilot study of 229 patients with CD.⁶²

Because multimodal models continue to improve, future AI-developed disease activity scores will combine multiple data sources such as imaging, endoscopy, and histology to produce a more nuanced assessment of disease activity and phenotypes (Fig. 1). Several crucial components are still needed to implement widespread, automated, population-level AI disease activity assessments. Substantial infrastructure is needed to digitize pathology whole-slide images, imaging studies, and store endoscopy video recordings. AI stands ready to not only standardize disease activity scoring in IBD health care delivery, but to be the tool necessary to develop new scoring systems capable of capturing a more comprehensive and precise assessment of disease burden, ultimately enhancing patient management, comparative effectiveness, and the evaluation of future therapeutics.

Key takeaways:

- AI-driven assessment tools for endoscopy, histology, and imaging in IBD show strong potential to standardize disease activity scoring, improve clinical trial efficiency, and reduce reliance on subjective and resource-intensive human evaluation.
- In endoscopy and histology, AI models demonstrate accuracy comparable to experts using established scoring

systems such as MES and Geboes, with added capabilities such as predicting disease flares.

- Advanced imaging applications, including CT, MRI, IUS, and VCE, along with novel AI-generated scoring systems and multimodal models, are expanding objective, granular, and predictive disease assessments, paving the way for improved patient management.

AI IN IBD PRIORITY: DECREASING THE TIME FROM THERAPEUTIC DISCOVERY TO REGULATORY APPROVAL WITH AI INTEGRATION

AI has the potential to transform multiple aspects of IBD therapeutic development, not only by providing AI-powered standard measures and potentially decreased costs, as discussed earlier, but also by automating subject enrollment and reducing the need for placebo group participants.⁶³ Trial Generative Pre-trained Transformer (TrialGPT) is a readily programmable LLM that searches through diverse types of medical records to find potential trial candidates. Applied to oncology trials, TrialGPT had an AUC of 0.817 for correctly classifying trial subjects. Although originally evaluated in oncology, TrialGPT offers a promising model that could be adapted to IBD trial recruitment pending disease-specific validation and could be readily applied to IBD trials.⁶⁴ In addition, LLM-powered chatbots have been successfully used to monitor patient progress and response to intervention in smoking cessation randomized controlled trials (RCTs).⁶⁵ IBD drug development trials could adapt these chatbot-based patient monitoring methods not only in industry-sponsored drug development studies but also scale them to population-level disease monitoring studies. Together, these examples show promise for leveraging AI-assistance to reduce the labor and expense associated with traditional clinical studies.

An additional challenge in clinical trials is declining recruitment because of concerns about being placed in a placebo arm.¹⁸ Ideally, future RCTs will address this by leveraging real-time adaptive trial designs with continuous AI-based monitoring to reduce placebo exposure.^{66,67} Digital twins are an excellent example of generative AI capabilities. Digital twins produce synthetic patients based on historical features that may be able to replace a large portion of placebo patients needed in clinical trials. Using data from a clinical trial in acute myeloid leukemia, we matched synthetic digital twin placebo patients with the demographic and genetic components of original placebo subjects and had similar simulated responses and outcomes than actual placebo subjects.⁶⁸

Postmarketing surveillance for adverse events (AEs) associated with IBD TIM therapy currently relies on prospective registry studies, spontaneous postmarketing reporting (eg, the Food and Drug Administrations [FDA] AE Reporting System), and/or analyses of billing code

data.⁶⁹⁻⁷¹ To address the underreporting and suboptimal sensitivity and specificity of current AE surveillance methods, LLM models have been developed to automate the detection of treatment-emergent AEs recorded in outpatient IBD clinic notes.⁴ These tools can help identify new safety signals and facilitate more informed discussions with patients about treatment risks. Overall, AI has the potential to improve the entire process of medication research and development from long-term safety monitoring.

Key takeaways:

- AI can enhance IBD therapeutic development by automating trial recruitment (eg, with models such as TrialGPT), reducing placebo use through digital twins, and supporting real-time adaptive trial designs to boost participation and efficiency.
- LLM-powered tools, including chatbots, show potential for scalable patient monitoring during clinical trials and real-world studies, reducing labor and cost burdens.
- AI-driven postmarketing surveillance using LLMs applied to clinical notes may significantly improve detection of AEs, strengthening long-term drug safety monitoring.

AI IN IBD PRIORITY: OPTIMIZE THERAPEUTIC SELECTION WITH AI-DRIVEN PRECISION MEDICINE TOOLS

The complex interplay of genetic, dietary, environmental, and microbial factors in the pathogenesis of IBD is increasingly being characterized.^{72,73} However, current analytical tools remain limited in addressing fundamental research questions about the origins of the disease. Groups have used AI as a method to generate hypotheses behind these complex disease mechanisms. For example, AI has been used to investigate new genetic signatures of UC and how these relate to the MES.⁷⁴ In CD, AI has been used to analyze the molecular profiles of CD enteroids compared with controls, identifying fatty acid metabolism as an important change in CD.⁷⁵ By integrating multiomics data, clinical records, and real-world evidence, AI can help identify novel disease subtypes and therapeutic targets. AI is being applied in preclinical research to accelerate the identification of promising therapeutic molecules and predict drug toxicity.⁷⁶⁻⁷⁸ Notably, a generative AI-developed, gut-restricted prolyl hydroxylase inhibitor (ISM012-042) for UC demonstrated efficacy in resolving colitis and restoring intestinal barrier function in a phase Ib trial.⁷⁹ The AI-developed molecule, ISM012-042, demonstrates the potential of AI to increase the therapeutics under investigation.

In addition to detangling the mechanisms behind the disease, AI will be an important tool in improving our understanding of the biology behind treatment response to reduce uncontrolled disease. For example, Verstockt et al⁶ used AI to identify multiple genes in colon tissue associated with endoscopic remission after vedolizumab

treatment but not anti-tumor necrosis factor therapy. AI may also be able to clarify which patients are unlikely to experience disease progression and could potentially be monitored off therapy, as shown in the work of Cushing et al,⁸⁰ who used AI to analyze whole transcriptome profiles to isolate features associated with indolent CD. IBD care stands to benefit from the potential of AI-driven tools capable of generating highly personalized insights to predict therapeutic response and risk based on multidimensional profiles, paving the way for precision medicine and improved clinical outcomes.

Key takeaways:

- AI has shown promise in its ability to advance our understanding of IBD pathogenesis by integrating multi-omics, clinical data, and real-world evidence to uncover novel disease mechanisms, subtypes, and therapeutic targets.
- In both UC and CD, AI has identified new genetic and molecular signatures and has even contributed to the development of novel therapeutics.
- AI models aim to improve the prediction of treatment response and disease progression, with the hope of enabling more personalized and precise IBD management through biologically driven patient stratification.

AI IN IBD PRIORITY: IMPROVE IBD ADMINISTRATIVE WORKFLOWS WITH AI

The chronic and complex nature of IBD necessitates detailed and comprehensive clinical documentation to effectively communicate with a multidisciplinary care team—an effort that demands significant time from clinicians.⁸¹ In addition, payors increasingly require extensive documentation to approve TIM therapies essential for many patients with IBD. To alleviate administrative burden, multiple companies are developing generative AI and NLP solutions designed to streamline clinical documentation. These technologies help reduce clinician workload by automating the review of historical clinical records and generating structured clinical notes. Many of these AI-driven tools use audio from patient-clinician encounters to transcribe conversations into prespecified note templates, some of which also produce summaries in nontechnical language for patients.⁸²⁻⁸⁹ Beyond documentation, AI is also being applied to enhance billing accuracy and minimize insurance denials by ensuring that clinical notes contain the necessary billing codes and justification for treatment.⁹⁰⁻⁹² AI-enabled tools are also being integrated to track prior authorization status automatically, reducing manual follow-ups.

Major U.S. insurance networks, including UnitedHealth, Humana, and CVS, have already implemented AI to automate claims processing and prior authorization. Notably, before AI-driven prior authorization, 73% of denials were overturned on appeal, underscoring the administrative inefficiencies of the system.⁹³ However, concerns have

emerged regarding transparency in AI-driven insurance processing because automated denials may lead to unnecessary delays and potential patient harm.^{94,95} These concerns have sparked class-action lawsuits against insurers using AI for automated denials.⁹⁶ Ultimately, it is hoped that future AI tools will be able to collate and analyze information to provide the necessary supporting documentation to minimize denials and the significant administrative burden associated with IBD care, leading to improved health outcomes for patients.

For decades, internet search engines have played a key role in patient education and support.⁹⁷ Now, generative AI is enabling patients to input their health information and receive personalized medical insights with unprecedented ease. Major technology companies along with startups such as are likely to develop AI-driven, direct-to-consumer products to meet the growing demand for accessible health information and lifestyle coaching. However, the accuracy and consistency of these AI-generated outputs currently require additional validation. To ensure the reliable dissemination of high-quality medical information, further efforts are needed—presenting a valuable opportunity for clinician partnerships in AI development.

Key takeaways:

- AI-powered tools show promise in streamlining clinical documentation, prior authorization tracking, and billing in IBD care, significantly reducing clinician workload and improving administrative efficiency.
- Despite widespread adoption by major insurers, AI-driven claims processing and prior authorization systems have raised concerns about transparency and patient harm because of automated denials, prompting legal and ethical scrutiny.
- Generative AI is also transforming patient engagement by offering personalized health information, though clinician involvement is essential to ensure the accuracy, safety, and credibility of these emerging patient-facing tools.

AI IN IBD PRIORITY: LARGE PUBLIC DE-IDENTIFIED MULTIMODAL IBD DATA SETS TO MAXIMIZE THE POTENTIAL OF AI IN IBD

Many of the currently available AI initiatives in medicine are driven by publicly available high-quality data sets such as chest radiographs, retinal images, and intensive care data, which provide the opportunity to benchmark model performance. However, 2 important pieces are necessary to drive the potential integration and impact of AI in IBD. First, one of the most significant bottlenecks to the development of AI in IBD remains data scarcity. There are no large public de-identified data sets specifically designed for IBD. The future of innovation and growth of AI in IBD will require large-scale, multi-institutional, de-identified data sets, which include both structured data such as billing codes and laboratories and unstructured

data such as notes, procedure reports, pathology reports, and imaging. Importantly, de-identification protects patients from data breaches.⁹⁸⁻¹⁰⁰ There are multiple algorithmic approaches, both open source and commercial, to ensure all data are de-identified at scale.¹⁰¹ Currently, only clinical trial data are available in IBD on platforms such as Vivli and YODA.^{102,103} Collaborative frameworks are needed to improve standardized approaches for collecting, organizing, and sharing real-world IBD health care data, including initiatives such as the Observational Medical Outcomes Partnership (OMOP) Common Data Model and Fast Healthcare Interoperability Resources (FHIR).¹⁰⁴ Once data are harmonized into standardized schemas, federated learning can be used as an ML approach that enables AI models to be trained across multiple institutions or data sources without the need to share or centralize raw data. Instead of transferring sensitive patient data to a single location for training, federated learning enables each participating institution to train a shared AI model locally on their own data. Only model updates—such as learned parameters or gradients—are sent to a central aggregator, which combines these updates to improve the global model while keeping patient data secure and private. The American Society for Gastrointestinal Endoscopy AI Institute is endeavoring to create a de-identified multi-institutional endoscopy video data set, Endonet, to address the challenge of data scarcity. An excellent example of AI's potential to reduce variability in IBD care is its role in dysplasia detection using comprehensive endoscopic and histologic databases. By enhancing endoscopists' ability to identify the often subtle features of dysplasia and predict histologic outcomes, AI could help reduce missed lesions while also minimizing unnecessary biopsies.¹⁰⁵

An area of investigation addressing data scarcity is developing varying degrees of synthetic data. Synthetic data include tabular, text-based, or time-series data generated by AI models mirroring an original data set's characteristics.¹⁰⁶ Synthetic data hold the promise of potentially accelerating the research pipeline of AI in IBD; however, models deployed in clinical settings should be evaluated, and if needed, fine-tuned, on original data.¹⁰⁷ Synthetic data offer approaches to enlarge data sets that are too small, potentially reducing bias, improving fairness, and enhancing privacy protection.¹⁰⁸ Given that the quality of AI model output is directly related to the quality of training data, significant additional research is needed to investigate the potential value of synthetic data in IBD to better understand the opportunities and limitations of this approach. Second, the development of IBD AI should prioritize aligning technological innovation with clearly defined clinical use-cases, rather than focusing on the computational techniques using only available data.¹⁰⁹ By addressing data scarcity and aligning AI innovation with meaningful clinical applications, the future of AI in IBD

can be shaped to maximize its impact on patient care, research, and health care delivery.

Key takeaways:

- The advancement of AI in IBD is limited by the absence of large, public, de-identified data sets, underscoring the urgent need for multi-institutional data sharing initiatives and standardized frameworks such as OMOP and FHIR.
- Innovative solutions such as federated learning and synthetic data generation offer promising strategies to overcome data scarcity while preserving patient privacy and improving model robustness and equity.
- To ensure meaningful impact, future AI development in IBD must align with clearly defined clinical use-cases, focusing on applications that directly enhance patient care, research, and health care delivery.

AI IN IBD PRIORITY: REGULATION TO SUPPORT INTEGRATION OF AI INTO IBD CLINICAL PRACTICE

Deploying clinical AI in IBD will require rigorous oversight. There is currently no FDA-approved AI for IBD. Similar to investigational therapies undergoing phased regulatory approval, we suggest a similar incremental approach to the approval of clinical AI in IBD, as shown in [Figure 2](#).¹⁰⁹ In addition, to minimize bias and systematic error, regulatory guidance specifying requirements for data set size, disease activity distribution, demographics, and adequate representation of underrepresented patient groups will be needed to ensure high confidence in model development and optimal performance. Local institution governance will be necessary to continuously monitor for changes in data patterns that can affect model performance.¹¹⁰

Not only will new methods of regulation and local governance be necessary for clinical AI in IBD, but advancing the emerging field of research into clinician–AI interaction will also be fundamental to its success. AI methods have been criticized for limited insight into their inner workings, with many calling for increased model explainability.¹¹¹ However, interpretability is not the only key component for optimal human–AI outcomes as seen in the RCT by Jabbour et al.¹¹² When clinicians were shown a biased AI model, their performance declined by over 11% despite being explicitly informed that the AI was using nonrelevant information. Model explainability did not fully mitigate the harmful effects of the biased AI on clinician performance. Education for health care professionals, delivered through current training programs and specialty societies, will be essential to support the effective use of AI.

For clinical AI solutions to be effectively adopted in IBD care, they must seamlessly integrate into existing software workflows, necessitating interoperability with electronic

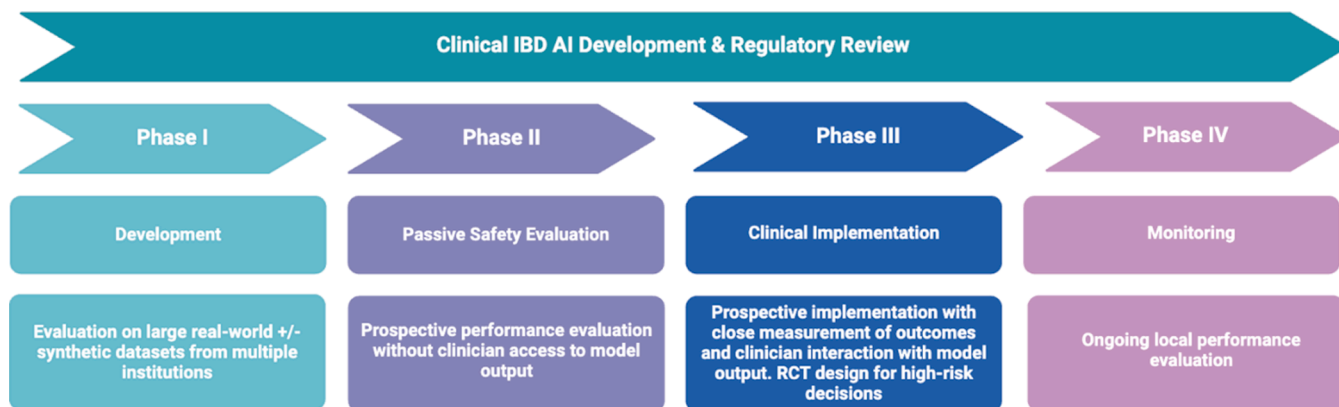


Figure 2. Proposed clinical AI in the IBD development pathway. *RCT*, Randomized controlled trial; *AI*, artificial intelligence; *IBD*, inflammatory bowel disease. Adapted from Patel et al.¹⁰⁹

health records and other clinical systems. In addition the development of appropriate billing codes or reimbursement frameworks will be essential to support implementation within health care payment models.

Successfully integrating AI into IBD management will require not only robust regulatory frameworks and continuous oversight but also reimbursement frameworks and a deep understanding of how clinicians interact with AI to ensure these tools enhance, rather than hinder, clinical decision-making.¹¹³

Key Takeaways:

- The safe deployment of clinical AI in IBD will require a regulatory approval process with guidance on data quality, representativeness, and ongoing institutional oversight to monitor model performance.
- Effective integration also depends on advancing research into clinician-AI interactions, as model explainability alone may not prevent harm from biased AI, highlighting the need for robust education and training for health care professionals.
- Ultimately, successful adoption will rely on a combination of regulatory rigor, local governance, and clinician preparedness to ensure AI tools enhance decision-making and improve patient care in IBD.

CONCLUSIONS AND FUTURE DIRECTIONS

The future of IBD care and research stands to benefit from the responsible and innovative integration of AI. To fully realize AI's potential in IBD, we must invest in robust de-identified public data sets, federated learning frameworks, ethical AI deployment, and drive forward ground-breaking research that transforms IBD management. This will require collaboration among clinicians, researchers, technology developers, health care institutions, policy-makers, and patient advocates. By working together, we can develop AI-driven solutions that enhance diagnostics, optimize treatment strategies, streamline clinical work-

flows, and improve patient outcomes. We call on all stakeholders to unite in shaping a future where AI is a tool for accelerating research and delivering high-quality, personalized, and equitable care for individuals with IBD. The time to act is now—let's build this future together.

DISCLOSURE

The following authors disclosed financial relationships: R.W. Stidham: funding support from AbbVie, Bristol Myers Squibb, Johnson and Johnson Innovative Medicine, and Merck; consultant or on advisory boards for AbbVie, Bristol Myers Squibb, CorEvitas, Eli Lilly, Gilead, Janssen, and Oruka Therapeutics; holds intellectual property on cross-sectional imaging and endoscopic analysis technologies licensed by the University of Michigan to AMI, llc, PreNovo, llc, and EIQ, llc with an equity interest in PathwaysGI, Inc. S. Parasa: Consultant at Fujifilm USA, Sebel Pharmaceuticals, Truveta, Allen Institute for Artificial Intelligence, Seattle; equity interest in DeepmediQ, Kayah.AI LLC; research support from Microsoft research and Google cloud. P. Sharma: Consultant at Olympus Corporation; Boston Scientific; Cipla; Medtronic; Takeda; Samsung Biopis; CDx; Exact Science; grant support from Fujifilm Erbe Medical; Ironwood Pharmaceuticals. U. Mahadevan: Consultant at AbbVie, Abivax, Bristol Myers Squibb, Boehringer Ingelheim, Celltrion, Enveda, Gilead, Janssen, Lilly, Pfizer, Protagonist, Roivant, Takeda, Trex; Data and Safety Monitoring Board: Merck. R. W. Stidham has funding support for this work from NIDDK (DK124779). U. Mahadevan has funding support from the Leona and Harry Helmsley Charitable Trust. The other author disclosed no financial relationships.

ACKNOWLEDGMENTS

Figures created using [BioRender.com](https://www.biorender.com).

- www.giejournal.org

- laser endomicroscopy in inflammatory bowel diseases. *Gastrointest Endosc* 2019;89:626-36.
39. Udristoiu AL, Stefanescu D, Gruionu G, et al. Deep learning algorithm for the confirmation of mucosal healing in Crohn's disease, based on confocal laser endomicroscopy images. *J Gastrointest Liver Dis* 2021;30:59-65.
40. Bossuyt P, Nakase H, Vermeire S, et al. Automatic, computer-aided determination of endoscopic and histological inflammation in patients with mild to moderate ulcerative colitis based on red density. *Gut* 2020;69:1778-86.
41. Stidham RW, Enchakalody B, Waljee AK, et al. Assessing small bowel stricturing and morphology in Crohn's disease using semi-automated image analysis. *Inflamm Bowel Dis* 2020;26:734-42.
42. Lamash Y, Kurugol S, Freiman M, et al. Curved planar reformatting and convolutional neural network-based segmentation of the small bowel for visualization and quantitative assessment of pediatric Crohn's disease from MRI. *J Magn Reson Imaging* 2019;49:1565-76.
43. Puylaert CAJ, Schüffler PJ, Naziroglu RE, et al. Semiautomatic assessment of the terminal ileum and colon in patients with Crohn disease using MRI (the VIGOR++ Project). *Acad Radiol* 2018;25:1038-45.
44. Majtner T, Brodersen JB, Herp J, et al. A deep learning framework for autonomous detection and classification of Crohn's disease lesions in the small bowel and colon with capsule endoscopy. *Endosc Int Open* 2021;9:E1361-70.
45. Klang E, Barash Y, Margalit RY, et al. Deep learning algorithms for automated detection of Crohn's disease ulcers by video capsule endoscopy. *Gastrointest Endosc* 2020;91:606-13.e2.
46. Ferreira JPS, de Mascarenhas Saraiva MJDQE, Afonso JPL, et al. Identification of ulcers and erosions by the novel Pillcam™ Crohn's capsule using a convolutional neural network: a multicentre pilot study. *J Crohns Colitis* 2022;16:169-72.
47. Higuchi N, Hiraga H, Sasaki Y, et al. Automated evaluation of colon capsule endoscopic severity of ulcerative colitis using ResNet50. *PLoS One* 2022;17:e0269728.
48. Meng J, Luo Z, Chen Z, et al. Intestinal fibrosis classification in patients with Crohn's disease using CT enterography-based deep learning: comparisons with radiomics and radiologists. *Eur Radiol* 2022;32: 8692-705.
49. Holland R, Patel U, Lung P, et al. Automatic detection of bowel disease with residual networks. *Predictive Intelligence in Medicine: Second International Workshop, PRIME 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings 2*. Springer; 2019:151-9.
50. Dolinger MT, Calabrese E, Pizzolante F, et al. Current and novel uses of intestinal ultrasound in inflammatory bowel disease. *Gastroenterol Hepatol* 2023;19:447.
51. Kucharzik T, Wilkens R, D'agostino M-A, et al. Early ultrasound response and progressive transmural remission after treatment with ustekinumab in Crohn's disease. *Clin Gastroenterol Hepatol* 2023;21:153-63.e12.
52. de Voogd F, van Wassenae EA, Mookhoek A, et al. Intestinal ultrasound is accurate to determine endoscopic response and remission in patients with moderate to severe ulcerative colitis: a longitudinal prospective cohort study. *Gastroenterology* 2022;163:1569-81.
53. Carter D, Albshesh A, Shimon C, et al. Automatized detection of Crohn's disease in intestinal ultrasound using convolutional neural network. *Inflamm Bowel Dis* 2023;29:1901-6.
54. Gu P, Chang J-H, Carter D, et al. Radiomics-based analysis of intestinal ultrasound images for inflammatory bowel disease: a feasibility study. *Crohns Colitis* 2024;6:otae034.
55. Baloesu C, Bailitz J, Cheema B, et al. Artificial intelligence-guided lung ultrasound by nonexperts. *JAMA Cardiol* 2025;10:245-53.
56. Barash Y, Azaria L, Soffer S, et al. Ulcer severity grading in video capsule images of patients with Crohn's disease: an ordinal neural network solution. *Gastrointest Endosc* 2021;93:187-92.
57. Cardoso P, Mascarenhas M, Afonso J, et al. Deep learning and minimally invasive inflammatory activity assessment: a proof-of-concept study for development and score correlation of a panendoscopy convolutional network. *Therap Adv Gastroenterol* 2024;17:17562848241251569.
58. Brodersen JB, Jensen MD, Leenhardt R, et al. Artificial intelligence-assisted analysis of pan-enteric capsule endoscopy in patients with suspected Crohn's Disease: a study on diagnostic performance. *J Crohns Colitis* 2024;18:75-81.
59. Li X, Liang D, Meng J, et al. Development and validation of a novel computed-tomography enterography radiomic approach for characterization of intestinal fibrosis in Crohn's disease. *Gastroenterology* 2021;160:2303-16.e11.
60. Meng J, Luo Z, Chen Z, et al. Intestinal fibrosis classification in patients with Crohn's disease using CT enterography-based deep learning: comparisons with radiomics and radiologists. *Eur Radiol* 2022;32:8692-705.
61. Chirra P, Sleiman J, Gandhi NS, et al. Radiomics to detect inflammation and fibrosis on magnetic resonance enterography in stricturing Crohn's disease. *J Crohns Colitis* 2024;18:1660-71.
62. Stidham RW, Enchakalody B, Wang SC, et al. Artificial intelligence for quantifying cumulative small bowel disease severity on CT-enterography in Crohn's disease. *Am J Gastroenterol* 2022;10:14309.
63. Sedano R, Solitano V, Vuyyuru SK, et al. Artificial intelligence to revolutionize IBD clinical trials: a comprehensive review. *Therap Adv Gastroenterol* 2025;18:17562848251321915.
64. Rybinski M, Kusa W, Karimi S, et al. Learning to match patients to clinical trials using large language models. *J Biomed Inform* 2024;159:104734.
65. Bricker JB, Sullivan BM, Mull KE, et al. Efficacy of a conversational chatbot for cigarette smoking cessation: protocol of the QuitBot full-scale randomized controlled trial. *Contemp Clin Trials* 2024;147:107727.
66. Rudrapatna VA, Cheng Y-W, Feuille C, et al. Creation of an ustekinumab external control arm for Crohn's disease using electronic health records data: a pilot study. *PLoS One* 2023;18:e0282267.
67. Bordukova M, Makarov N, Rodriguez-Esteban R, et al. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials. *Expert Opin Drug Discov* 2024;19:33-42.
68. Eckardt J-N, Hahn W, Röllig C, et al. Mimicking clinical trials with synthetic acute myeloid leukemia patients using generative artificial intelligence. *NPJ Digit Med* 2024;7:76.
69. FDA. Questions and Answers on FDA's Adverse Event Reporting System (FAERS). Available at: <https://www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers#:~:text=The%20FDA%20Adverse%20Event%20Reporting,that%20were%20submitted%20to%20FDA>. Accessed June 1, 2025.
70. Thein D, Egeberg A, Skov L, et al. Absolute and relative risk of new-onset psoriasis associated with tumor necrosis factor- α inhibitor treatment in patients with immune-mediated inflammatory diseases: a Danish nationwide cohort study. *JAMA Dermatol* 2022;158:997-1004.
71. Chaparro M, Garre A, Ricart E, et al. Short and long-term effectiveness and safety of vedolizumab in inflammatory bowel disease: results from the ENEIDA registry. *Aliment Pharmacol Ther* 2018;48:839-51.
72. Pizarro TT, Stappenbeck TS, Rieder F, et al. Challenges in IBD research: preclinical human IBD mechanisms. *Inflamm Bowel Dis* 2019;25(Supplement_2):S5-12.
73. Syed AH, Abujabal HAS, Ahmad S, et al. Advances in inflammatory bowel disease diagnostics: machine learning and genomic profiling reveal key biomarkers for early detection. *Diagnostics (Basel)* 2024;14:1182.
74. Yang Y, Hua Y, Zheng H, et al. Biomarkers prediction and immune landscape in ulcerative colitis: findings based on bioinformatics and machine learning. *Comput Biol Med* 2024;168:107778.
75. Fernandes P, Sharma Y, Zulqarnain F, et al. Identifying metabolic shifts in Crohn's disease using 'omics-driven contextualized computational metabolic network models. *Sci Rep* 2023;13:203.
76. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596:583-9.
77. Lin Z, Akin H, Rao R, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 2023;379:1123-30.

78. Basile AO, Yahi A, Tatonetti NP. Artificial intelligence for drug toxicity and safety. *Trends Pharmacol Sci* 2019;40:624-35.
79. Fu Y, Ding X, Zhang M, et al. Intestinal mucosal barrier repair and immune regulation with an AI-developed gut-restricted PHD inhibitor. *Nat Biotechnol* 2024;43:1-6.
80. Cushing KC, Mclean R, McDonald KG, et al. Predicting risk of postoperative disease recurrence in Crohn's disease: patients with indolent Crohn's disease have distinct whole transcriptome profiles at the time of first surgery. *Inflamm Bowel Dis* 2019;25:180-93.
81. Sinsky C, Colligan L, Li L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. *Ann Intern Med* 2016;165:753-60.
82. Abridge. Available at: <https://www.abridge.com/>. Accessed June 1, 2025.
83. DAX Copilot. Available at: <https://www.nuance.com/healthcare/dragon-ai-clinical-solutions/dax-copilot.html>. Accessed June 1, 2025.
84. DeepScribe. Available at: <https://www.deepscribe.ai/>. Accessed June 1, 2025.
85. AWS HealthScribe. Available at: <https://aws.amazon.com/healthscribe/>. Accessed June 1, 2025.
86. AutoScribe. Available at: <https://www.mutuohealth.com/>. Accessed June 1, 2025.
87. Code Care. Available at: <https://codecare.ai/>. Accessed June 1, 2025.
88. Freed. Available at: <https://www.getfreed.ai/>. Accessed June 1, 2025.
89. Clinical Notes. Available at: <https://www.clinicalnotes.ai/>. Accessed June 1, 2025.
90. Suki. Available at: <https://www.suki.ai/>. Accessed June 1, 2025.
91. Eleos Compliance. Available at: <https://eleos.health/compliance/>. Accessed June 1, 2025.
92. Charta. Available at: <https://www.chartahealth.com/>. Accessed June 1, 2025.
93. Office of Inspector General. Some Medicare Part D Beneficiaries Face Avoidable Extra Steps That Can Delay or Prevent Access to Prescribed Drugs. U.S. Department of Health and Human Services; 2019. Report OEI-09-16-00411. Available at: <https://oig.hhs.gov/reports/all/2019/some-medicare-part-d-beneficiaries-face-avoidable-extra-steps-that-can-delay-or-prevent-access-to-prescribed-drugs>. Accessed August 27, 2025.
94. Constant BD, Elder A, Long MD, et al. Insurer medication utilization barriers are associated with decreased insurance satisfaction and adverse clinical outcomes: an IBD partners survey. *Am J Gastroenterol* 2022;10:14309.
95. Shah NB, Zuckerman AD, Hosteng KR, et al. Insurance approval delay of biologic therapy dose escalation associated with disease activity in patients with inflammatory bowel disease. *Dig Dis Sci* 2023;68:4331-8.
96. Congress, Senate, Committee on Homeland Security and Governmental Affairs. Refusal of Recovery: How Medicare Advantage Insurers Have Denied Patients Access to Post-Acute Care: Majority Staff Report. 2024. Available at: https://www.govinfo.gov/app/details/GOVPUB-Y4_G74_9-PURL-gpo234149. Accessed August 27, 2025.
97. Mandl KD. How AI could reshape health care-rise in direct-to-consumer models. *JAMA* 2025;333:1667-9.
98. Ford E, Pillinger S, Stewart R, et al. What is the patient re-identification risk from using de-identified clinical free text data for health research? *AI Ethics* 2025;5:1-14.
99. Cloud Healthcare API. Google cloud data de-identification. Available at: <https://cloud.google.com/healthcare-api/docs/concepts/de-identification>. Accessed June 1, 2025.
100. Apexon. Apexon data de-identification. Available at: <https://www.apexon.com/>. Accessed June 1, 2025.
101. Norgeot B, Muenzen K, Peterson TA, et al. Protected Health Information filter (Philter): accurately and securely de-identifying free-text clinical notes. *NPJ Digit Med* 2020;3:57.
102. YODA. YODA project. Available at: <https://yoda.yale.edu/>. Accessed June 1, 2025.
103. Vivli. Available at: <https://vivli.org/>. Accessed June 1, 2025.
104. Mandl KD, Gottlieb D, Mandel JC. Integration of AI in healthcare requires an interoperable digital data ecosystem. *Nat Med* 2024;30:631-4.
105. Vinsard DG, Fetzter JR, Agrawal U, et al. Development of an artificial intelligence tool for detecting colorectal lesions in inflammatory bowel disease. *iGIE* 2023;2:91-101.e6.
106. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. *NPJ Digit Med* 2023;6:186.
107. Jordon J, Szpruch L, Houssiau F, et al. Synthetic Data—what, why and how? *arXiv:2205.03257*.
108. Guillaudeux M, Rousseau O, Petot J, et al. Patient-centric synthetic data generation, no reason to risk re-identification in biomedical data analysis. *NPJ Digit Med* 2023;6:37.
109. Patel MR, Balu S, Pencina MJ. Translating AI for the clinician. *JAMA* 2024;332:1701-2.
110. Nong P, Hamasha R, Singh K, et al. How academic medical centers govern AI prediction tools in the context of uncertainty and evolving regulation. *Massachusetts Medical Society*; 2024:Alp2300048.
111. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat Mach Intell* 2019;1:206-15.
112. Jabbour S, Fouhey D, Shepard S, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: a randomized clinical vignette survey study. *JAMA* 2023;330:2275-84.
113. Henry KE, Kornfield R, Sridharan A, et al. Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *NPJ Digit Med* 2022;5:97.

Abbreviations: AEs, adverse events; AI, artificial intelligence; AUC, area under the curve; CD, Crohn's disease; CDS, cumulative disease score; CV, computer vision; DL, deep learning; FDA, Food and Drug Administration; FHIR, Fast Healthcare Interoperability Resources; GHAS, Global Histology Activity Score; IBDs, inflammatory bowel diseases; IUS, intestinal US; LLMs, large language models; MES, Mayo Endoscopic Score; ML, machine learning; NLP, natural language processing; OMOP, Observational Medical Outcomes Partnerships; RCTs, randomized controlled trials; TIM, targeted immune modulator; TrialGPT, Trial Generative Pre-trained Transformer; UC, ulcerative colitis; VCE, video capsule endoscopy.

*On behalf of the American Society for Gastrointestinal Endoscopy Artificial Intelligence Institute.

Copyright © 2025 by the American Society for Gastrointestinal Endoscopy 0016-5107/\$36.00
<https://doi.org/10.1016/j.gie.2025.12.248>

Received December 3, 2025. Accepted December 3, 2025.

Current affiliations: Department of Medicine, The Dr. Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, New York, New York, USA (1), Division of Gastroenterology and Hepatology, Department of Internal Medicine, Mayo Clinic, Scottsdale, Arizona, USA (2), Division of Gastroenterology, Department of Internal Medicine, Michigan Medicine, Ann Arbor, Michigan, USA (3), Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA (4), Michigan Institute for Data Science, University of Michigan, Ann Arbor, Michigan, USA (5), Department of Gastroenterology, Swedish Medical Center, Seattle, Washington, USA (6), Department of Gastroenterology, Kansas City Veterans Affairs Medical Center, Kansas City, Missouri, USA (7), Division of Gastroenterology and Hepatology, Department of Internal Medicine, University of Kansas School of Medicine, Kansas City, Kansas, USA (8), Division of Gastroenterology, University of California, San Francisco, San Francisco, California, USA (9).

Corresponding author: Anna L. Silverman, MD, Department of Medicine, The Dr. Henry D. Janowitz Division of Gastroenterology, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place, New York, New York, 10029, USA. E-mail: Anna.Silverman@mountsinai.org.