




# Artificial intelligence to revolutionize IBD clinical trials: a comprehensive review

Rocio Sedano , Virginia Solitano , Sudheer K. Vuyyuru, Yuhong Yuan, Jurij Hanžel, Christopher Ma , Olga Maria Nardone and Vipul Jairath

*Ther Adv Gastroenterol*

2025, Vol. 18: 1–20

DOI: 10.1177/  
17562848251321915

© The Author(s), 2025.  
Article reuse guidelines:  
sagepub.com/journals-  
permissions

**Abstract:** Integrating artificial intelligence (AI) into clinical trials for inflammatory bowel disease (IBD) has potential to be transformative to the field. This article explores how AI-driven technologies, including machine learning (ML), natural language processing, and predictive analytics, have the potential to enhance important aspects of IBD trials—from patient recruitment and trial design to data analysis and personalized treatment strategies. As AI advances, it has potential to improve long-standing challenges in trial efficiency, accuracy, and personalization with the goal of accelerating the discovery of novel therapies and improve outcomes for people living with IBD. AI can streamline multiple trial phases, from target identification and patient recruitment to data analysis and monitoring. By integrating multi-omics data, electronic health records, and imaging repositories, AI can uncover molecular targets and personalize trial strategies, ultimately expediting drug development. However, the adoption of AI in IBD clinical trials encounters significant challenges. These include technical barriers in data integration, ethical concerns regarding patient privacy, and regulatory issues related to AI validation standards. Additionally, AI models risk producing biased outcomes if training datasets lack diversity, potentially impacting underrepresented populations in clinical trials. Addressing these limitations requires standardized data formats, interdisciplinary collaboration, and robust ethical frameworks to ensure inclusivity and accuracy. Continued partnerships among clinicians, researchers, data scientists, and regulators will be essential to establish transparent, patient-centered AI frameworks. By overcoming these obstacles, AI has the potential to enhance the efficiency, equity, and efficacy of IBD clinical trials, ultimately benefiting patient care.

Correspondence to:

**Vipul Jairath**  
Division of  
Gastroenterology,  
Department of Medicine,  
Western University,  
London, ON, Canada

Department of  
Epidemiology and  
Biostatistics, Western  
University, London, ON,  
Canada

Lawson Health Research  
Institute, Room A10-219,  
University Hospital, 339  
Windermere Rd, London,  
ON N6A 5A5, Canada  
[vjairath@uwo.ca](mailto:vjairath@uwo.ca)

**Rocio Sedano**  
Division of  
Gastroenterology,  
Department of Medicine,  
Western University,  
London, ON, Canada

Department of  
Epidemiology and  
Biostatistics, Western  
University, London, ON,  
Canada

Lawson Health Research  
Institute, London, ON,  
Canada

**Virginia Solitano**  
Department of  
Epidemiology and  
Biostatistics, Western  
University, London, ON,  
Canada

Division of  
Gastroenterology  
and Gastrointestinal  
Endoscopy, IRCCS  
Ospedale San Raffaele,  
Università Vita-Salute San  
Raffaele, Milan, Lombardy,  
Italy

**Sudheer K. Vuyyuru**  
Division of  
Gastroenterology,  
Department of Medicine,  
Western University,  
London, ON, Canada

## Plain language summary

### Artificial intelligence in IBD clinical trials

Inflammatory Bowel Disease (IBD), including Crohn's disease and ulcerative colitis, poses significant challenges for clinical trials, such as difficulties in recruiting participants, variations in disease presentation, and inconsistent treatment responses. Artificial intelligence (AI) is increasingly recognized as a solution to these challenges, improving recruitment, data analysis, personalized care, and trial design. AI can enhance recruitment by analyzing medical records to match patients to trials efficiently. AI tools can automate this process, improving both efficiency and diversity. Additionally, AI can predict dropout risks, helping researchers plan better and maintain trial integrity. IBD trials generate complex datasets that require advanced analysis. AI can process these large datasets to identify patterns in disease progression and treatment efficacy, also improving the accuracy of endoscopic and histological assessments, providing deeper insights into the disease. AI can enable personalized treatments by predicting responses based on genetics, biomarkers, and medical history. Real-time monitoring through wearable devices supports early interventions, improving patient outcomes and disease management. Adaptive trial designs might also benefit from AI, allowing protocols to adjust based on interim results.

**Yuhong Yuan**

Division of  
Gastroenterology,  
Department of Medicine,  
Western University,  
London, ON, Canada

Lawson Health Research  
Institute, London, ON,  
Canada

**Jurij Hanžel**

Department of  
Gastroenterology,  
University Medical Centre  
Ljubljana, University  
of Ljubljana, Ljubljana,  
Slovenia

**Christopher Ma**

Department of Community  
Health Sciences, Cumming  
School of Medicine,  
University of Calgary,  
Calgary, AB, Canada

Division of  
Gastroenterology and  
Hepatology, Department  
of Medicine, University  
of Calgary, Calgary, AB,  
Canada

**Olga Maria Nardone**

Gastroenterology,  
Department of Public  
Health, University Federico  
II of Naples, Naples, Italy

This enhances trial efficiency, ethical standards, and participant safety, while ensuring accurate data collection. However, implementing AI requires addressing data privacy, fairness, and regulatory compliance. Transparent, secure, and inclusive AI models are essential to build trust and ensure equitable benefits across all patient populations. AI is transforming IBD clinical trials by streamlining recruitment, improving data analysis, personalizing care, and optimizing trial design. By addressing challenges proactively, we can unlock AI's full potential, leading to more efficient trials and better outcomes for patients.

**Keywords:** artificial intelligence (AI), clinical trials, inflammatory bowel disease, machine learning (ML), patient recruitment

Received: 2 December 2024; revised manuscript accepted: 4 February 2025.

## Introduction

Inflammatory bowel disease (IBD), encompassing Crohn's disease (CD) and ulcerative colitis (UC), is a chronic and relapsing condition that presents significant challenges for patient management and in conducting clinical research in an efficient manner.<sup>1,2</sup> Recruitment rates for pharmaceutical IBD clinical trial continue to decline with average global recruitment rates of 0.1 patient/center/trial. The reasons for poor recruitment are complex and multifactorial and include factors related to patient characteristics, overengineered protocols, use of placebo and increasing availability of commercial agents. In addition, heterogeneity in disease presentation, variability in treatment response, and the requirement for long-term management necessitate innovative approaches in clinical trials. Traditional methodologies have often failed to address these complexities, highlighting the need for advanced technologies to provide more precise and efficient solutions at all stages of a clinical trial including patient recruitment, endpoint assessment, and trial monitoring.

In recent years, artificial intelligence (AI) and machine learning (ML) have emerged as powerful medical tools, offering unprecedented data analysis, pattern recognition, and predictive modeling capabilities. These technologies have the potential to enhance aspects of IBD clinical trials and may provide solutions to challenges that have historically impeded progress in this field.<sup>3</sup> However, the promise of AI can only be fully realized if these technologies are scientifically rigorous, clinically valid, and aligned with regulatory

standards to ensure patient safety and improve outcomes.<sup>4,5</sup>

This review explores the potential impact that AI could have on IBD clinical trials, focusing on key areas such as patient recruitment, data analysis, personalized medicine, and trial design. Additionally, we discuss the ethical, regulatory, and practical considerations that must be addressed to ensure the responsible integration of AI in clinical trials for IBD.

A summary of definitions of commonly used terminology is included in Table 1.

## AI in patient recruitment: Streamlining and enhancing recruitment for IBD clinical trials

Recruiting patients for IBD clinical trials remains a significant challenge.<sup>14</sup> Despite some advances, issues such as overestimating eligible populations, limited patient awareness, logistical difficulties, and competition for participants still persist.<sup>15</sup> Traditional recruitment approaches are often resource-heavy and struggle to efficiently identify and enroll appropriate candidates. AI offers promising solutions by analyzing large datasets, including electronic health records (EHRs) and patient-reported outcomes. These technologies may help to more efficiently and accurately match eligible participants with trial criteria, improving recruitment criteria.<sup>16,17</sup>

## Screening activities

AI-driven tools like natural language processing (NLP) are improving clinical trial recruitment by

**Table 1.** Summary of definitions of commonly used terminology.

Term	Abbreviation	Definition	Ref
Artificial intelligence	AI	A field of computer science that involves the simulation of human intelligence in machines, enabling them to perform tasks that typically require human cognition, such as learning, reasoning, and decision-making.	6
Machine learning	ML	A subtype of AI that focuses on developing algorithms that allow computers to learn from and make decisions based on data without being explicitly programmed for specific tasks.	6,7
Deep learning	DL	A subtype of ML that uses neural networks with multiple layers (hence “deep”) to analyze large datasets and extract high-level features for tasks such as image and speech recognition.	6,7
Neural networks	NN	Computational models inspired by the human brain that are composed of interconnected nodes (neurons) and are used in ML and DL to identify patterns and make predictions.	7
Natural language processing	NLP	A branch of AI that enables computers to understand, interpret, and generate human language, often used for tasks like text analysis, language translation, and chatbot interactions.	8
Large language models	LLMs	A type of AI model trained on vast amounts of text data to understand and generate human-like language, often used for tasks like summarization, translation, and conversational AI.	9,10
Shapley additive explanations	SHAP	A framework used in ML to explain the output of predictive models by assigning importance values to each input feature, helping users understand how a model arrives at its predictions.	11
Explainable artificial intelligence	XAI	AI systems designed to provide clear, understandable explanations for their predictions and decisions, improving trust and accountability in their applications.	12
Artificial intelligence fairness 360	AIF360	Toolkit developed by IBM to detect and mitigate bias in AI models, ensuring equitable and inclusive outcomes in AI-driven decision-making processes, such as clinical trial recruitment.	13

analyzing clinical notes and unstructured data to identify potential participants that might be missed by traditional screening methods. Leveraging NLP and ML, generative AI can automate the evaluation of eligibility criteria against medical histories, drastically reducing the need for manual reviews. This allows researchers and clinical staff to potentially pre-screen hundreds of candidates in just minutes, speeding up pre-screening activities. One recent advancement in this field is TrialGPT, a model designed to improve patient-trial matching.<sup>18</sup> Using large language models, TrialGPT analyzes patient

medical records and compares them with trial eligibility criteria. Trained on data from 184 patients with complex conditions—predominantly cancer and other chronic diseases such as cardiovascular disease, diabetes, and rare genetic conditions—and 18,238 annotated clinical trials, TrialGPT not only determines patient suitability but also provides detailed explanations for its decision.<sup>18</sup> When tested on a larger dataset of patients across oncology and various chronic disease populations, TrialGPT demonstrated strong performance. Its explanations aligned closely with those of human experts, effectively ranking trials and

excluding those for which patients were ineligible. However, some errors were noted due to limitations in the underlying language models.

Generative AI, using tools such as chatbots and virtual assistants, can also reduce the screening burden on clinical trial sites by handling initial participant screening and communication. AI-enabled platforms such as myTrialsConnect enhance participant interactions, gather trial-specific data, and even schedule appointments, thus improving accessibility and workload management.<sup>19</sup>

#### *Enhancing diversity in enrollment*

In IBD clinical trials, recruitment bias can lead to the underrepresentation of minority populations, which impacts the generalizability of findings. AI Fairness 360 (AIF360) developed by IBM Research might tackle this issue by ensuring that AI-driven recruitment algorithms do not disproportionately exclude these groups, fostering more inclusive and equitable study populations.<sup>13</sup> By addressing bias, the toolkit allows researchers to increase the likelihood of fair and representative samples that reflect the diversity of the population affected by the disease being studied, and to ensure that eligibility criteria are applied equitably to all groups, avoiding bias that could exclude minority populations.

Furthermore, AI can predict patient dropout rates and adherence to trial protocols, enabling proactive management of these issues, which is crucial for maintaining trial integrity.<sup>20,21</sup>

#### **Enhancing data analysis through AI: Utilizing ML to analyze complex IBD trial outcome data**

IBD clinical trials generate large and complex datasets, including clinical, imaging, biomarker data, and genomic data. AI, particularly ML and deep learning algorithms, can process high-dimensional data to uncover hidden patterns or correlations critically for understanding disease progression, treatment responses, and patient subgroups.<sup>22,23</sup>

#### *EHR data analysis using ML methods*

A key application of AI in IBD research involves leveraging ML techniques to analyze EHR-derived

data. These methods allow for integrating patient demographics, physiological measurements, disease history, clinical questionnaires, histology, serum biomarkers, and drug exposure to uncover insights that traditional analyses may overlook. ML-based models, such as XGBoost and deep learning approaches, can identify complex, non-linear relationships that influence disease progression and therapeutic outcomes.

#### *Predicting response to therapy*

A recent study by Harun et al. shed light on the role of AI in shaping future clinical trial designs, particularly by identifying stratification factors that can optimize treatment effectiveness and improve patient outcomes.<sup>24</sup> The authors conducted a post hoc analysis of four randomized controlled trials (RCTs) of etrolizumab in patients with UC, using advanced ML techniques to assess which patient factors impact remission. XGBoost ML models were used to evaluate the effect of various patient-level data on the likelihood of achieving remission. To interpret the complex predictions, the SHAP (SHapley Additive exPlanations) framework clarified which factors were most influential. The data analyzed included demographics, physiological measurements, disease history, clinical questionnaires, histology, serum biomarkers, and drug exposure. The models performed well, achieving an area under the receiver operating characteristic curve (AUROC) of  $0.74 \pm 0.03$  for induction and  $0.75 \pm 0.06$  for maintenance. By using AI techniques, the study was able to analyze a large, complex dataset and reveal nonlinear relationships and interactions that traditional methods might miss. The use of XGBoost improved the predictive accuracy of remission based on diverse variables, offering deeper insights into patient outcomes. The SHAP framework further enhanced understanding by identifying key factors influencing remission, aiding in patient stratification and optimizing treatment strategies for future trials.

#### *Endoscopy*

Endoscopic assessment is the cornerstone to establish patient eligibility for IBD trial participation and to estimate the efficacy of trial interventions. Blinded central endoscopic reading is the current standard which, compared to local endoscopic reading, increases objectivity, minimizes

variability, reduces placebo rates, and consequently maximizes effect sizes.<sup>25–29</sup> Nonetheless, even agreement between expert central readers is imperfect,<sup>25,30</sup> and several scoring conventions were added to existing evaluative indices with the purpose of harmonizing scores.<sup>31</sup> Disagreement between central readers further complicates the scoring process by introducing the need for outcome adjudication, where multiple read algorithms to resolve disagreement and assign a final score are possible, each with their advantages and disadvantages.<sup>32,33</sup>

Several AI models have been developed to deliver reliable and accurate readings of endoscopic videos in UC. In a recently published meta-analysis, 12 studies were included, with 9 studies evaluating the Mayo endoscopic score (MES)<sup>34</sup> and 3 the Ulcerative Colitis Endoscopic Index of Severity (UCEIS)<sup>35</sup> as the reference standard.<sup>36</sup> Overall, the sensitivity and specificity of AI for endoscopic assessment was high; both for still images (sensitivity 91%, specificity 89%) and videos (sensitivity 86%, specificity 91%). A notable finding was the high heterogeneity between studies with  $I^2$  values exceeding 90%.

Several aspects of study design should however be considered to correctly contextualize these findings and identify future research priorities. All studies used a human expert reader as the reference standard. As the training of convolutional neural networks, the AI tool used in endoscopy, depends on the human reader reference, they, by definition, cannot yet surpass human performance in terms of accuracy, whereas gains in efficiency and throughput may be considerable. A further potential use of AI is using it by default to screen all endoscopies, performed at a given site with the purpose of identifying patients with endoscopically active disease (MES 2 or 3) and flagging them for potential inclusion in trials. Studies included in the meta-analysis all used dichotomized outcomes, for example, MES 0–1 versus 2–3. Whilst endoscopic remission was defined as a MES 0–1 in the past,<sup>37</sup> the two scores are no longer conflated in contemporary trials: a score of 0 denotes endoscopic remission and a score of 1 endoscopic improvement, reflecting two different outcomes.<sup>38</sup> The distinction between a score of 2 and 3 is also not insignificant as baseline endoscopic activity may serve as a stratification factor for randomization. It should be

acknowledged that only three of the studies supported their models with external validation cohorts.<sup>39–41</sup> Finally, the high heterogeneity persisted even in sensitivity analyses separating studies based on still image versus video assessment and based on the numbers of images evaluated. The high variability between studies could therefore be the result of discrepancies in image annotation, image pre-processing, and training algorithms.<sup>42</sup> Developing AI algorithms for endoscopic assessment of CD remains an unmet research need, studies thus far have focused on video capsule endoscopy, which does not feature in regulatory clinical trials.

Using AI-based algorithms in clinical trials has been shown to be feasible as a recurrent neural network model performed favorably compared to human readers for the evaluation of full-length endoscopy videos in a phase II trial of mirikizumab.<sup>43</sup> It should be noted, however, that this trial utilized a single central reader paradigm, and it remains unknown how to best integrate AI algorithms in multiple central reader paradigms which require outcome adjudication. Currently, we are lacking studies to inform optimum positioning of AI-based algorithms in reading paradigms: it is unknown, whether the algorithm should replace the local reader, the central reader, or perhaps even both readers. An often-cited limitation of the MES is the fact that it defaults to the worst affected area of the colon visualized, regardless of potential changes in disease extent.<sup>44</sup> An AI-based solution integrating both endoscopic disease severity and disease extent is the cumulative disease score (CDS).<sup>45</sup> A notable advantage of this system, tested on the ustekinumab trial dataset is its superior ability to discriminate between the ustekinumab arm and the placebo arm—a simulated sample size calculation indicated that 50% fewer patients would be needed to demonstrate a difference with CDS compared to the MES. Although the latter remains the regulatory standard, CDS could be used in early drug development programs to detect between-arm differences with smaller numbers of patients. AI could conceivably recognize endoscopic lesions and patterns, which are not part of established endoscopic indices and therefore be potentially more sensitive to change. This could be particularly helpful in early drug development to guide decisions whether to continue the clinical program.



### *Histology*

Histological remission is currently understood as an adjunct to endoscopic remission indicating a deeper level of healing.<sup>46</sup> It is not yet considered a treatment target, but its potential role is being evaluated in a randomized trial.<sup>47</sup> Similarly to endoscopy, histological assessment depends on scoring indices, which face challenges similar to those of endoscopic indices: inter-rater reliability is imperfect, scoring can be time-consuming and requires expertise.<sup>48</sup> In UC histology, AI models have been used to develop novel scoring indices,<sup>49,50</sup> to replace human readers for established indices,<sup>51,52</sup> and to evaluate individual histological features, such as the presence of eosinophils<sup>53</sup> and basal cell plasmacytosis.<sup>54</sup>

Overall, models developed to evaluate biopsies using existing indices have shown encouraging sensitivity and specificity to detect histological remission.<sup>51,52</sup> One of the systems was developed using clinical trial data, demonstrating feasibility in this setting.<sup>51</sup> Analogously to AI models for endoscopic assessment, currently developed algorithms predict histological remission as a binary outcome, but cannot provide grading of inflammatory activity. Arguably, this is less of a limitation for histology than it is for endoscopy as precise grading of histologically active disease is less relevant and has little impact on the interpretation of clinical trial results.

A further area of development of AI is also the deployment of algorithms to help guide human pathologists identify the main areas of interest within a given biopsy fragment.<sup>51</sup> Novel algorithms also promise to detect histological features beyond those included in established histological scoring indices, which could be more informative for predicting subsequent treatment outcomes. In a study of 114 patients with UC achieving endoscopic improvement ( $MES \leq 1$ ), a deep learning model successfully quantified the ratio between the goblet cell mucus area and epithelial cells, a lower ratio was associated with an increased rate of disease relapse within the subsequent 12 months.<sup>55</sup> Even more impressively, a ML-based algorithm was able to identify 18 histologic features, which were able to predict which patients with pediatric UC would not respond to treatment with mesalamine alone.<sup>56</sup> These features, discovered in an inception cohort of 292 patients, were later tested in an external validation cohort with almost identical performance (AUROC 0.89

in the development cohort and 0.88 in the validation cohort). More recently, Ohara et al.<sup>57</sup> developed an advanced AI system incorporating semantic segmentation and object detection models to identify neutrophils in hematoxylin and eosin-stained WSIs. This system not only detects neutrophils in the epithelium and lamina propria but also predicts components of the Nancy Histological Index and the PICaSSO Histologic Remission Index.<sup>41</sup> Notably, the AI-predicted histological scores correlated well with pathologists' assessments (Spearman's  $\rho = 0.68-0.80$ ;  $p < 0.05$ ).

In another study, Peyrin-Biroulet et al.<sup>58</sup> utilized automated image analysis combined with ML to evaluate histological disease activity based on the Nancy index in 200 histological images from UC patients. The AI system's performance was compared to assessments by four independent histopathologists. Despite limitations due to the small annotated dataset required for AI training,<sup>59</sup> the study reported high correlations both among histopathologists (89.33) and between the AI system and histopathologists (87.20).

### *Radiology*

Radiological assessment is likely to have an increasingly prominent role in clinical trials in IBD. Transmural healing is defined as an adjunct to endoscopic remission, reflecting a deeper level of healing in CD,<sup>46</sup> its potential advantage over contemporary treatment goals is under evaluation in an ongoing randomized trial (NCT06257706). Fibrostenosing CD is an area of unmet therapeutic need, and the development of potential anti-fibrotic agents is a research priority.<sup>60</sup> The recent publication of dedicated indices<sup>61,62</sup> extends the role of cross-sectional imaging beyond evaluating inflammation.

Detection and characterization of strictures is an area well-suited to AI models. Good concordance has been shown between (semi)-automated measurements and expert radiologist assessment for key elements, such as bowel wall thickness, pre-stenotic dilation, and minimum luminal diameter.<sup>63,64</sup> Notably, AI was able to quantify intestinal fibrosis with an AUROC exceeding 0.800 compared to the reference standard of histopathological assessment of the resection specimen.<sup>64</sup> Automated assessment was non-inferior to expert radiological assessment and considerably faster.

AI has not yet been evaluated for radiological assessment of perianal fistulizing CD, which is expected to be quite challenging given the complex morphology and heterogeneity of this disease phenotype.

Emerging research also indicates that convolutional neural networks are able to accurately identify abnormal bowel wall thickening on images obtained with intestinal ultrasound, although this remains to be proven in studies with larger sample sizes and also using cine loops as opposed to still images.<sup>65,66</sup>

A summary of endoscopic, histologic and radiology AI modalities can be found in Table 2.

#### *Synthesizing multimodal clinical trial data*

Possibly the greatest opportunity for harnessing the power of AI in clinical trials lies in the analysis of complex multimodal data to predict outcomes. Existing endoscopic indices have limited predictive capability for subsequent disease evolution and histological indices focus, perhaps unduly, on neutrophils. An algorithm quantifying red pixels in endoscopy videos of UC, the red-density index, performed acceptably for predicting 5-year clinical remission.<sup>69,70</sup> An algorithm based on endoscopy, supported by endocytoscopy, classified patients by risk of clinical relapse in real-time.<sup>71</sup> The AI-based PICaSSO Histologic Remission Index successfully estimated the likelihood of a flare of UC at 1 year.<sup>52</sup> AI approaches are also well suited to the analysis of complex proteomic and microbiomic data. An AI algorithm supported classification of patients based on the relative abundance of 92 inflammatory protein, which was associated with subsequent response to treatment with infliximab.<sup>67</sup> A neural network algorithm integrating clinical and microbiome data, demonstrated an acceptable predictive capability for clinical remission after 14 weeks of treatment with vedolizumab in CD.<sup>68</sup> The use of wearable devices, such as smart watches, for monitoring IBD and predicting subsequent disease flares is under active investigation. The abundance of clinical and biomarker data gathered through these devices could optimally be analyzed using AI-based methods to develop novel digital outcomes and predict future disease evolution.<sup>72,73</sup>

Additionally, AI could integrate multi-omics data to identify novel biomarkers, which can serve as

surrogate endpoints in clinical trials, thereby accelerating drug development.<sup>74</sup> AI-driven analytics could also enhance the understanding of patient heterogeneity in IBD, enabling the identification of distinct disease subtypes. This stratification informs the development of targeted therapies, ultimately leading to more personalized treatment approaches.<sup>75</sup>

#### **Personalized medicine and AI: Tailoring treatment strategies to individual patients**

The current management of IBD is challenging, as the disease varies widely in severity and response to treatment among patients. With the availability of several classes of advanced therapies, choosing a suitable therapeutic agent which would result in optimal response is challenging. Currently there are no tools that can accurately predict response to any given agent. Traditional treatment protocols often involve the best clinical judgment based on available evidence, clinical records, social factors, and local institutional policies. Personalized medicine is an emerging approach that moves away from a one-size-fits-all paradigm in healthcare, instead tailoring medical treatment to the individual characteristics of each patient. For chronic, heterogeneous conditions such as IBD, this approach holds significant potential. Patients with IBD differ widely in their genetic profiles, disease severity, and response to medications.

The advent of AI presents a revolutionary opportunity to optimize and personalize treatment strategies for IBD, tailoring care to individual patient characteristics. AI can analyze large, complex datasets that include genetic information, clinical records, environmental factors, and treatment responses, enabling clinicians to make more precise and effective decisions in IBD management.<sup>76,77</sup> This process significantly reduces the time spent on trial-and-error treatments, improves clinical outcomes, and minimizes the risk of adverse effects from ineffective therapies.

AI models can continuously learn and adapt based on new data, enabling real-time personalization of treatment strategies.<sup>78</sup> These adaptive strategies could be particularly beneficial in managing IBD, where disease activity can fluctuate over time.<sup>79–81</sup> AI can assist physicians in several stages of management of patients with IBD including choosing appropriate agent early at the

**Table 2.** Summary of endoscopic, histologic, and radiology AI modalities.

Modality	Advantages	Disadvantages	Examples of use
Endoscopy	Reliable and accurate AI scoring of endoscopic disease activity in ulcerative colitis compared to human readers Potentially more efficient than human central readers	High heterogeneity between studies Not available for Crohn's disease Unknown optimal reading paradigm (local reader vs central reader vs AI) Most algorithms provide only binary outcomes (remission vs no remission)	Use of the Cumulative Disease Score in the ustekinumab development program—appears more sensitive in detecting differences between the drug arm and placebo arm, potentially requiring smaller sample sizes <sup>45</sup> An AI model performed well on full-length trial videos from the mirikizumab program—both for the Mayo score and the Ulcerative Colitis Endoscopic Index of Severity <sup>43</sup>
Histology	Good concordance between AI scoring and human pathologists Supporting human reading by highlighting regions of interest Potential for discovering novel histological features, which are not part of established indices, but are potentially associated with subsequent outcomes	Algorithms provide only binary outcomes (remission vs no remission)	Feasibility demonstrated on trial datasets for both Crohn's disease and ulcerative colitis <sup>51</sup>
Radiology	Encouraging initial results for characterizing strictures in Crohn's disease, compared to human readers Potential for use in intestinal ultrasound and perianal fistulizing Crohn's disease	Limited research available	Good concordance between AI algorithms and human readers to characterize strictures <sup>63,64</sup>
Multimodal trial data	Potential to overcome limitations of traditional analytical approaches to better predict treatment outcomes Potential to analyze large biomarkers datasets and data from wearable devices	Algorithmic bias	Predicting response to treatment with infliximab based on an array of 92 proteins <sup>67</sup> Integrating clinical and microbiome data to predict clinical remission with vedolizumab in Crohn's disease <sup>68</sup>

AI, artificial intelligence.

time of diagnosis, predicting disease progression and exacerbations allowing early intervention.

*AI in choosing appropriate therapeutic agent*

There are several factors that determine response to a given agent. AI has been shown to be useful in predicting treatment response to various drugs in cancer therapy and antibacterial therapy.<sup>82,83</sup> Similarly, AI can predict which patients are more likely to respond to certain biologic therapies based on their genetic and microbiome composition.<sup>77,84</sup> Several studies demonstrated predictive ability of AI models in predicting response to

various advanced therapies such as anti-TNFs, vedolizumab, and ustekinumab in patients with CD.<sup>85–88</sup> Some of these studies used clinical and laboratory data, while others used genotype data. Recent research has identified hundreds of genetic loci associated with IBD, yet these genetic insights have not been fully integrated into clinical practice due to their complexity. AI can synthesize this genomic information with other patient-specific data to create predictive models that anticipate how a patient's disease will progress and how they will respond to different treatments. Therefore, the use of AI models can potentially predict best treatment option for a given patient.



In the coming years, AI-based histopathological studies are expected to make significant contributions to IBD management. Preliminary data<sup>89</sup> showed that computational pathology algorithms can identify cytokines, such as IL-23 signaling activity, from H&E images. This could significantly enhance our understanding of disease path mechanisms and optimize treatment options for patients with IBD.

#### *AI in predicting disease progression*

The unpredictable course of IBD is one of the most challenging aspects of managing the disease. Patients often alternate between periods of active inflammation and remission, with some experiencing frequent complications, such as fistulas, strictures, or the need for surgery. Predicting when a patient will experience a disease exacerbation or develop complications is essential for timely intervention and disease management. ML algorithms can be trained on large datasets that include clinical records, laboratory results, imaging studies, and lifestyle factors to identify patterns and predictors of disease progression.<sup>90</sup> These algorithms can then generate risk profiles for individual patients, estimating the likelihood of a flare-up, complication, or need for surgery. For example, AI can assess inflammatory biomarkers, such as C-reactive protein or fecal calprotectin, alongside clinical symptoms and patient-reported outcomes, to predict when a patient is at high risk of a non-response to therapy.<sup>91</sup> Such predictive models allow physicians to modify treatments proactively, such as increasing medication doses or initiating new therapies before the patient experiences a relapse. For example, in a study from Korea, the ML model for prediction of IBD-related outcomes at 5 years after diagnosis yielded an area under the curve of 0.86 (95% CI: 0.82–0.92). This model performed consistently across a range of other datasets, enabling physicians to perform close follow-up based on the patient's risk level.<sup>92</sup> A novel ML model based on data of 20,368 veteran health administration patients substantially improved ability to predict future IBD-related hospitalization and steroid use.<sup>93</sup> Furthermore, AI can predict long-term outcomes in IBD patients, guiding decisions about the intensity of treatment. For example, patients at high risk of developing complications might benefit from early, aggressive therapy with biologics or immunosuppressants, while those with a lower risk profile could be managed with

less intensive treatments. This individualized approach to disease management can reduce overtreatment, minimize side effects, and improve patient quality of life. Recently, Wang *et al.*<sup>94</sup> developed a deep learning framework aimed at predicting postoperative recurrence in CD. The model automatically analyzed the muscular layer and myenteric plexus, integrating clinical data to evaluate myenteric plexitis severity and recurrence risk. This approach sheds light on the mechanisms underlying postoperative recurrence and offers potential for enhancing long-term disease management.

Wearable devices, mobile health applications, and home-based diagnostic tools can collect continuous data on a patient's symptoms, biometrics, and lifestyle factors. AI can analyze these data streams in real time, identifying subtle changes that may indicate an impending flare-up or treatment failure. For instance, fluctuations in inflammatory markers detected through home-based stool tests or blood samples can signal worsening disease activity. By integrating this data with AI algorithms, clinicians can be alerted to intervene early, preventing a full-scale relapse or the need for hospitalization. AI-powered applications can provide patients with personalized feedback based on their symptoms, response to questionnaires. These applications can remind patients to take their medications, track their symptoms, and alert them to seek medical attention if necessary. AI can further enhance these platforms by analyzing patterns in patient-reported outcomes, detecting early warning signs of non-adherence or treatment failure, and suggesting adjustments to the treatment plan. Real-time monitoring supported by AI not only improves disease management but also empowers patients to take a more active role in their care. This proactive approach has the potential to reduce the burden of IBD, both in terms of physical symptoms and the psychological toll of living with a chronic illness.

#### *Utilizing multi-omics data and AI to aid personalized medicine*

In recent years omics data analysis has gained importance and has helped in understanding pathogenesis of IBD. The integration of multi-omics and clinical data has been enhanced, leading to breakthroughs in disease diagnosis, drug discovery, and precision medicine. ML-based methods offer significant advantages in handling

large-scale datasets and can reveal patterns among a high number of features that traditional methods may fail to identify. Multi-omics analysis can help in holistically understanding pathogenesis, simultaneous changes in microbiome, biological processes which in turn can help the researchers in identifying potential targets. For instance, in a study by Lloyd-Price et al. extensive multi-omics molecular profiling was performed on 132 IBD patients.<sup>95</sup> The authors observed significant alterations in microbiota composition and function based on disease activity states. In another study, remission-associated multi-omic profiles were unique to each therapeutic class.<sup>96</sup> Recently integration of endoscopic, histological data with multi-omics has also been proposed.<sup>97</sup> Moreover, AI models offer the opportunity to identify a pioneering gut barrier-protective agents for IBD and forecasts the potential success of candidate agents in phase III trials. Sahoo et al. developed an AI-assisted approach for target identification and validation. This ML path has demonstrated the ability to predict epithelial barrier-related genes, such as PRKAB1, the  $\beta$ 1 subunit of the metabolic master regulator, AMPK, which might represent a novel target for gut barrier-protective therapies.<sup>98</sup>

However, AI application in multi-omics is still in infancy and major challenges of multi-omics AI models include the lack of generalization when applied to independent validation cohorts. The primary limitation of the clinical applicability of AI lies precisely in the high heterogeneity of the disease and its variations over time, leading to inadequate reproducibility and generalizability of predictive results and a possible overestimation of prediction accuracy. In the near future, the AI approaches are expected to be especially valuable in classifying already diagnosed patients into disease sub-phenotypes, predicting disease progression, and evaluating response to treatment.

#### **AI in trial design and monitoring: Enhancing adaptive trial designs and real-time participant monitoring**

The FDA has highlighted the importance of integrating AI and ML into drug and biological product development, particularly in clinical trial designs. Their discussion paper highlights the potential of AI to streamline the development process by enhancing the design and execution of clinical trials through adaptive methodologies,

real-time monitoring, and predictive modeling.<sup>4,5,99</sup> This framework not only enhances the efficiency and accuracy of clinical trials but also aligns with regulatory requirements to ensure patient safety and data integrity.<sup>100</sup>

One particular advance is the integration of AI for advancing adaptive trial designs, which allow for modifications based on interim data, and have the potential to improve trial efficiency and ethical integrity by minimizing the probability of being randomized to potentially less effective or safe therapies.<sup>101</sup> AI-driven adaptive trials can lead to more flexible studies by enabling real-time adjustments based on patient responses and emerging data.<sup>102–104</sup> For instance, whereas traditional frequentist trial designs have pre-specified endpoints where formalized hypothesis testing is carried out to evaluate efficacy and/or safety, AI-driven adaptive designs can potentially identify earlier signals of treatment efficacy or safety, enabling researchers to adjust dosing regimens, modify inclusion criteria, or in the most extreme situation, terminate trials early if necessary.<sup>105</sup> Multiple applications of AI in adaptive trial designs can be implemented. For example, AI-driven predictive analytics or simulation can be used for unbiased interim data analysis, AI-driven outcome prediction modeling can be used for sample size estimation (or re-estimation), AI-driven ML models can inform covariate- or response-adaptive randomization processes, and AI-driven models can be used to generate valid external control arms to help reduce the likelihood of placebo randomization.

In addition to trial design, AI may also facilitate more efficient and actionable real-time monitoring of participants to ensure safety in trial participation. Wearable devices and mobile apps integrated with AI can continuously collect and analyze patient data, providing insights into adherence, disease progression, and adverse events. This real-time monitoring could be essential for proactively managing patient safety and trial integrity.<sup>106,107</sup>

Moreover, AI has the potential to improve risk-based monitoring (RBM). Historically, RBM has involved a multifaceted approach for identifying, assessing, monitoring, and subsequently mitigating risks that pose threats to quality or safety in an RCT. Given the tremendous amount of data collected within a clinical trial, AI-driven data

monitoring systems can detect subtle changes in patient or site-level data that might indicate an adverse event, lack of efficacy, protocol deviations, or potentially site-related concerns that can prompt timely interventions. These systems have the potential to integrate much more data than centralized human monitors and can also help manage large-scale, decentralized trials by coordinating data from multiple sites and ensuring consistency in trial conduct.<sup>107,108</sup>

### **Challenges and ethical considerations: Addressing ethical, regulatory, and practical challenges**

The integration of AI in IBD clinical trials offers transformative possibilities but also presents significant ethical, regulatory, and practical challenges that must be addressed proactively. One primary ethical concern relates to data privacy, especially given the sensitive nature of health data used in AI models, which includes genetic, biomarker, and multi-omic information. To protect patient confidentiality while enabling AI models to function effectively, robust data security measures such as encryption, anonymization, and strict access protocols must be rigorously implemented. These measures should comply with regulatory standards like General Data Protection Regulation (GDPR) or the Health Insurance Portability and Accountability Act (HIPAA), which are essential for safeguarding patient information throughout AI/ML-driven clinical trials.<sup>109</sup>

Transparency and interpretability of AI algorithms are other critical ethical priorities. Many AI models, particularly deep learning systems, operate as “black boxes,” making predictions without easily explainable reasoning.<sup>110</sup> This lack of interpretability can hinder clinicians’ trust in AI-driven recommendations, posing risks of automated decisions that may not align with the best interests of patients. Ensuring transparency and accountability in AI decision-making processes is essential.<sup>111,112</sup> To address this, ongoing research and model development should focus on explainable AI frameworks and interpretable ML techniques, which can help clinicians understand and evaluate how models arrive at their conclusions, thereby enhancing clinical justification and maintaining clinician accountability.<sup>113,114</sup>

The development of attention-based models and post hoc explainability methods such as SHAP and Local Interpretable Model-Agnostic Explanations (LIME) are active areas of research aimed at improving the transparency of AI-driven decisions in IBD trials.<sup>115</sup>

Algorithmic bias is another significant concern. AI models trained on historically imbalanced datasets may inadvertently perpetuate existing biases, leading to the underrepresentation or misrepresentation of specific patient groups, such as racial or ethnic minorities, in clinical trials. Studies have shown that IBD trials have historically underrepresented minorities, which can limit the model’s effectiveness in identifying eligible participants from diverse racial or ethnic backgrounds.<sup>116–118</sup> Furthermore, this bias can also affect underrepresented subgroups of IBD, such as individuals with fistulizing CD, pouchitis, and extraintestinal manifestations, among other underserved conditions. This further marginalizes these populations and restricts our understanding of these complex and less common disease presentations. Addressing bias requires both methodological and policy-level interventions, such as diversifying training datasets and implementing bias-detection frameworks, adversarial debiasing methods, and fairness-aware ML algorithms.<sup>119,120</sup> These strategies can help ensure that AI recruitment algorithms equitably represent the diversity of the IBD patient population, promoting more inclusive and equitable clinical trials.<sup>116–118,121</sup> Foundation models, which leverage vast pre-trained datasets, have also been explored as a means to mitigate bias by improving generalizability and reducing the impact of imbalanced training samples.<sup>122</sup>

On the regulatory front, the FDA and other regulatory agencies have issued preliminary guidelines on the use of AI in clinical research; however, comprehensive regulations are still evolving.<sup>4,5</sup> The lack of standardized regulatory frameworks creates uncertainties for AI developers and researchers, especially concerning model validation, continuous learning protocols, and real-time adjustments in adaptive trials.<sup>123</sup> To navigate these complexities, ongoing collaboration among industry stakeholders, researchers, and regulatory agencies is crucial for establishing clear guidelines on the development, deployment, and monitoring of AI in clinical trials.<sup>124</sup> Regulatory bodies,

including the FDA, emphasize the importance of model validation and explainability, calling for regular algorithmic validation, retraining, and the establishment of processes to monitor updates and improvements in AI systems throughout the course of a trial.<sup>4,5,102</sup> The emergence of foundation models has further encouraged regulatory discussions, as their broad pre-training across multiple domains can reduce the frequency of retraining and improve transferability across diverse patient populations.<sup>125</sup>

From a practical standpoint, implementing AI-driven technologies in clinical trials requires substantial resources, including technical infrastructure, specialized personnel, and ongoing oversight to maintain model performance and data integrity.<sup>126</sup> For example, effective implementation in decentralized trials requires robust digital platforms for seamless data collection and integration across multiple sites.<sup>127</sup> Integrating AI models into clinical workflows also presents logistical challenges; clinicians and trial coordinators need training to understand, trust, and effectively utilize these tools. Practical guidelines for AI implementation should emphasize user-friendly interfaces, interoperability with existing systems, and consistent support to facilitate successful integration into clinical practice.<sup>4,126</sup>

Lastly, as AI increasingly plays a role in real-time patient monitoring through wearable devices and remote data collection,<sup>128,129</sup> patient autonomy and informed consent become critical ethical considerations.<sup>130</sup> Patients should be fully informed about how their data will be used, including the role of AI in monitoring their health and influencing treatment pathways. Informed consent processes must adapt to address AI-driven data analysis, providing clear assurances regarding data security, the purpose of data collection, and the limits of AI's predictive capabilities.<sup>131</sup> By prioritizing patient autonomy, transparency, and robust security measures, researchers can foster trust in AI-enabled IBD trials, thereby promoting equitable and ethical innovation in clinical research.

### Discussion

AI represents a paradigm shift in the landscape of IBD clinical trials and has a potential to expedite the drug development process, thereby making

safe and effective drugs available for the patient faster. AI can help at several stages of drug development process including but not limited to identification of potential molecular targets by integrating and analyzing multi-omics data, streamlining patient recruitment, enhancing data analysis, personalizing strategies, thereby optimizing trial design and monitoring<sup>107,132</sup> and post-marketing surveillance.

Realizing the potential of AI in clinical trials requires careful consideration of the ethical, regulatory, and practical challenges associated with its integration.<sup>123</sup> Despite its promise, AI technology is still in the early stages, and several obstacles must be overcome before its widespread implementation in drug development processes. Integration of data across various systems, including EHRs, laboratory databases, and imaging repositories, and integration of these databases can be a challenging task. Moreover, variability in data formats, terminologies, and quality across institutions can lead to inaccuracies in AI predictions. Additionally, privacy concerns and proprietary restrictions often limit access to the data needed to train and validate AI models.

Bias in AI models is another critical issue. If the datasets used to train AI systems are not representative of diverse populations, AI algorithms may produce biased predictions, leading to inaccurate outcomes.<sup>133,134</sup> The development of AI-driven tools for IBD must prioritize inclusivity, ensuring that these technologies benefit all patient populations, particularly those that have been historically underrepresented in clinical research.<sup>81,116</sup> By addressing these challenges, AI can fulfill its promise of transforming IBD clinical research, leading to more effective and equitable healthcare outcomes. For example, underrepresentation of minority groups or women could result in skewed trial results, while AI designs and implementations may unintentionally favor certain demographics, contributing to inequities in clinical trial recruitment and treatment recommendations. Furthermore, the use of AI systems requires careful attention to ethical considerations around sensitive personal data. Ensuring robust informed consent processes that address how AI technologies are employed is essential. Moreover, anonymizing and securing data while maintaining its utility for AI presents considerable challenges.



AI in clinical trials is still relatively novel, and regulatory frameworks remain under development. There is currently no universally accepted standard for validating AI tools in clinical trials, making it difficult to guarantee their reliability and reproducibility. Integrating AI into clinical trials necessitates significant adjustments to established workflows. Clinicians and trial staff often lack the technical expertise needed to effectively operate and interpret AI tools. This requires extensive training and collaboration with data scientists. Additionally, resistance to AI adoption may arise from stakeholders skeptical about the accuracy of these systems or concerned about job displacement. The development and implementation of AI systems in clinical trials can be resource-intensive, with high initial investment requirements. Running advanced AI algorithms demands substantial computing infrastructure, which may not be available in all clinical trial settings. Additionally, continuous updates, validation, and monitoring of AI systems are necessary to ensure their ongoing accuracy and relevance, contributing to long-term costs. AI models also require rigorous validation to ensure they can be applied effectively across diverse trial populations. However, many models are tested on limited datasets, raising concerns about their generalizability. Differences in data sources, trial protocols, and patient demographics can lead to inconsistent results when applying the same AI model across multiple settings.

While AI holds tremendous potential to transform clinical trials, its implementation faces numerous challenges. Overcoming these obstacles requires a multidisciplinary approach that involves collaboration between researchers, clinicians, data scientists, ethicists, and regulators. Key strategies include enhancing data standardization, ensuring transparency in AI systems, establishing clear regulatory guidelines, and fostering education and collaboration among stakeholders. Addressing these challenges will enable AI to become a powerful tool for making clinical trials more efficient, equitable, and effective. The FDA's guidance provides a comprehensive framework for the responsible use of AI in drug and biological product development, emphasizing the need for transparency, validation, and patient-centric approaches.<sup>4,5,135</sup> A recent report from World Health Organization also highlights the ethics and recommendations for governance of AI in health care.<sup>136</sup> As the field continues to evolve,

ongoing collaboration between researchers, regulators, and industry stakeholders will be essential to use AI's full potential in IBD research.

### Conclusion

The integration of AI into clinical trials for IBD represents a significant advancement in gastroenterology research and patient care ushering in a new era of precision medicine. AI technologies, including ML and predictive analytics, are revolutionizing trial design, patient recruitment, endpoint assessment, data analyses, personalized treatment strategies, and the monitoring and prediction of treatment responses. The ability to assess deeper levels of healing, such as barrier healing, will enhance therapeutic strategies and potentially organ-sparing approaches.

By leveraging large datasets, AI enhances the accuracy and diversity of participant selection while providing deeper insights into disease mechanisms. Its ability to customize treatment plans for individual patients promises improved outcomes and reduced side effects, which is crucial in IBD management due to the variability in patient responses. However, despite its potential, the adoption of AI presents critical challenges that require careful consideration. For instance, AI systems generated on supervised learning inherently depend on the assumption that the input data—often derived from physician diagnoses or clinical observations—is accurate. This reliance underscores the importance of high-quality, well-annotated datasets to minimize errors and biases. In the context of IBD, where diagnosis and disease characterization can be complex, ensuring the reliability of input data is essential to avoid perpetuating inaccuracies through AI-driven analyses.

Another important consideration is the cost associated with incorporating AI into clinical trials. Developing and maintaining AI systems requires significant investment in infrastructure, including high-performance computing capabilities, data integration platforms, and skilled personnel such as data scientists and bioinformaticians. Moreover, the ongoing need for algorithm validation, retraining, and compliance with regulatory standards adds to the financial burden. While these costs may be prohibitive for some institutions, they must be weighed against the potential long-term benefits, such as more efficient trials,



personalized treatments, and reduced healthcare expenditures resulting from improved disease management.

In conclusion, while AI represents a paradigm shift in IBD research and clinical trials, its successful implementation will depend on addressing these foundational challenges. Collaboration between researchers, clinicians, regulatory authorities, and industry stakeholders will be crucial to ensuring that AI technologies are accurate, transparent and accessible. By prioritizing data quality, cost-effectiveness, and ethical standards, the integration of AI has the potential to significantly enhance treatment outcomes and advance the IBD and gastroenterology field.

### Declarations

*Ethics approval and consent to participate*  
Not applicable.

*Consent for publication*  
Not applicable.

### Author contributions

**Rocio Sedano:** Conceptualization; Investigation; Methodology; Project administration; Writing – original draft; Writing – review & editing.

**Virginia Solitano:** Writing – original draft; Writing – review & editing.

**Sudheer K. Vuyyuru:** Writing – original draft; Writing – review & editing.

**Yuhong Yuan:** Methodology; Writing – review & editing.

**Jurij Hanžel:** Writing – original draft; Writing – review & editing.

**Christopher Ma:** Writing – original draft; Writing – review & editing.

**Olga Maria Nardone:** Writing – review & editing.

**Vipul Jairath:** Conceptualization; Supervision; Writing – review & editing.

*Acknowledgements*  
None.

### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

### Competing interests

R.S. has received consulting fees from AbbVie, Alimentiv, Inc., and Takeda. V.S. has received speaker's fees from Pfizer. S.K.V. has received consulting fees from Alimentiv. Y.Y. has no any competing interest. J.H. has received consulting fees from AbbVie, Alimentiv, Inc., Eli Lilly, and Takeda. C.M. has received consulting fees from AbbVie, Alimentiv, Amgen, AVIR Pharma, Inc., Bristol Myers Squibb, Celltrion, Eli Lilly, Ferring, Forte Biosciences, Fresenius Kabi, Gilead, Janssen, McKesson, Mirador Therapeutics, Mylan, Pendopharm, Pfizer, Prometheus Biosciences, Inc., Roche, Sanofi, Takeda, Tillotts Pharma; speaker's fees from AbbVie, Amgen, AVIR Pharma Inc, Alimentiv, Bristol Myers Squibb, Eli Lilly, Ferring, Fresenius Kabi, Janssen, Merck, Organon, Pendopharm, Pfizer, Sanofi, Takeda, Tillotts Pharma; royalties from Springer Publishing; research support from AbbVie, Eli Lilly, Ferring, Pfizer. O.M.N. has received speaker's fees from Janssen, Eli Lilly, Alfa Sigma, Ferring, Pfizer, Noös; Advisory Board from Nestlé. V.J. has received consulting/ advisory board fees from AbbVie, Alimentiv, Anaptyis Bio, Arena pharmaceuticals, Asahi Kasei Pharma, Asieris, Astra Zeneca, Bristol Myers Squibb, Celltrion, Eli Lilly, Endpoint Health, Entera, Ensho, Ferring, Flagship Pioneering, Fresenius Kabi, Galapagos, GlaxoSmithKline, Genentech, Gilead, Innomar, JAMP, Janssen, Merck, Metacrine, Mylan, MRM Health, Pandion, Pendopharm, Pfizer, Protagonist, Prometheus Biosciences, Reistone Biopharma, Roche, Roivant, Sandoz, Second Genome, Sorriso, Spyre, Synedgen, Takeda, Teva, Ventyx, Vividion; speaker's fees from, AbbVie, Ferring, Bristol Myers Squibb, Eli Lilly, Fresenius Kabi, Janssen, Pfizer, Shire, Takeda, and Tillotts.

### Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed.

### ORCID iDs

Rocio Sedano  <https://orcid.org/0000-0002-3311-638X>

Virginia Solitano  <https://orcid.org/0000-0002-2935-6778>

Christopher Ma  <https://orcid.org/0000-0002-4698-9948>

## References

1. Le Berre C, Honap S and Peyrin-Biroulet L. Ulcerative colitis. *Lancet* 2023; 402: 571–584.
2. Dolinger M, Torres J and Vermeire S. Crohn's disease. *Lancet* 2024; 403: 1177–1191.
3. Solitano V, Zilli A, Franchellucci G, et al. Artificial endoscopy and inflammatory bowel disease: welcome to the future. *J Clin Med* 2022; 11: 569.
4. FDA. U.S. Food and Drug Administration. Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan, vol. 2021. U.S. Food and Drug Administration, <https://www.fda.gov/media/145022/download> (2021, accessed 19 October 2024).
5. FDA. U.S. Food and Drug Administration. Using artificial intelligence & machine learning in the development of drug & biological products: discussion paper and request for feedback. FDA, <https://www.fda.gov/media/167973/download> (2021, accessed 19 October 2024).
6. Soori M, Arezoo B and Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, a review. *Cogn Robot* 2023; 3: 54–70.
7. Kufel J, Bargiel-Laczek K, Kocot S, et al. What is machine learning, artificial neural networks and deep learning?—examples of practical applications in medicine. *Diagnostics (Basel)* 2023; 13: 2582.
8. Soysal E, Wang J, Jiang M, et al. CLAMP—a toolkit for efficiently building customized clinical natural language processing pipelines. *J Am Med Inform Assoc* 2018; 25: 331–336.
9. Zhang K, Yang X, Wang Y, et al. Artificial intelligence in drug development. *Nat Med* 2025; 31: 45–59.
10. Ntinopoulos V, Rodriguez Cetina Biefer H, Tudorache I, et al. Large language models for data extraction from unstructured and semi-structured electronic health records: a multiple model performance evaluation. *BMJ Health Care Inform* 2025; 32: e101139.
11. Rodriguez-Perez R and Bajorath J. Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values. *J Med Chem* 2020; 63: 8761–8777.
12. Ali S, Abuhmed T, El-Sappagh S, et al. Explainable Artificial Intelligence (XAI): what we know and what is left to attain trustworthy artificial intelligence. *Inf Fusion* 2023; 99: 101805.
13. Bellamy RKE, Dey K, Hind M, et al. AI Fairness 360: an extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev* 2019; 63(4): 1–4:15.
14. Honap S, Jairath V, Danese S, et al. Navigating the complexities of drug development for inflammatory bowel disease. *Nat Rev Drug Discov* 2024; 23: 546–562.
15. Solitano V, Prins H, Archer M, et al. Toward patient centricity: why do patients with inflammatory bowel disease participate in pharmaceutical clinical trials? A mixed-methods exploration of study participants. *Crohns Colitis* 360 2024; 6: otae019.
16. Ahmad HA, East JE, Panaccione R, et al. Artificial intelligence in inflammatory bowel disease endoscopy: implications for clinical trials. *J Crohns Colitis* 2023; 17: 1342–1353.
17. Ismail A, Al-Zoubi T, El Naqa I, et al. The role of artificial intelligence in hastening time to recruitment in clinical trials. *BjR Open* 2023; 5: 20220023.
18. Jin Q, Wang Z, Floudas CS, et al. Matching patients to clinical trials with large language models. ArXiv 2024.
19. ClinicalTrialsArena.com. Elligo and Avallano launch AI-powered clinical trials platform, vol. 2024, <https://www.clinicaltrialsarena.com/news/elligo-and-avallano-launch-ai-powered-clinical-trials-platform/> (2023, accessed 20 October 2024).
20. Obermeyer Z and Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med* 2016; 375: 1216–1219.
21. Yin J, Ngiam KY and Teo HH. Role of artificial intelligence applications in real-life clinical practice: systematic review. *J Med Internet Res* 2021; 23: e25759.
22. Char DS, Shah NH and Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med* 2018; 378: 981–983.
23. Stafford IS, Gosink MM, Mossotto E, et al. A systematic review of artificial intelligence and machine learning applications to inflammatory

- bowel disease, with practical guidelines for interpretation. *Inflamm Bowel Dis* 2022; 28: 1573–1583.
24. Harun R, Lu J, Kassir N, et al. Machine learning-based quantification of patient factors impacting remission in patients with ulcerative colitis: insights from Etrolizumab Phase III Clinical Trials. *Clin Pharmacol Ther* 2024; 115: 815–824.
  25. Feagan BG, Sandborn WJ, D'Haens G, et al. The role of centralized reading of endoscopy in a randomized controlled trial of mesalamine for ulcerative colitis. *Gastroenterology* 2013; 145: 149–157 e2.
  26. Duijvestein M, Jeyarajah J, Guizzetti L, et al. Response to placebo, measured by endoscopic evaluation of Crohn's disease activity, in a pooled analysis of data from 5 randomized controlled induction trials. *Clin Gastroenterol Hepatol* 2020; 18: 1121–1132.e2.
  27. Almradi A, Sedano R, Hogan M, et al. Clinical, endoscopic, and safety placebo rates in induction and maintenance trials of Crohn's disease: meta-analysis of randomised controlled trials. *J Crohns Colitis* 2022; 16: 717–736.
  28. Sedano R, Hogan M, Nguyen TM, et al. Systematic review and meta-analysis: clinical, endoscopic, histological and safety placebo rates in induction and maintenance trials of ulcerative colitis. *J Crohns Colitis* 2022; 16: 224–243.
  29. Solitano V, Hogan M, Singh S, et al. Placebo rates in Crohn's disease randomized clinical trials: an individual patient data meta-analysis. *Gastroenterology* 2025; 168(2): 344–356.
  30. Khanna R, Zou G, D'Haens G, et al. Reliability among central readers in the evaluation of endoscopic findings from patients with Crohn's disease. *Gut* 2016; 65: 1119–1125.
  31. Khanna R, Ma C, Hogan M, et al. Standardizing scoring conventions for Crohn's disease endoscopy: an international RAND/UCLA Appropriateness Study. *Clin Gastroenterol Hepatol* 2023; 21: 2938–2950 e6.
  32. Reinisch W, Mishkin DS, Oh YS, et al. Impact of various central endoscopy reading models on treatment outcome in Crohn's disease using data from the randomized, controlled, exploratory cohort arm of the BERGAMOT trial. *Gastrointest Endosc* 2021; 93:174–182 e2.
  33. Jairath V, Zou GY, Ma C, et al. Routine incorporation of the local read in Crohn's disease clinical trials? Not so fast. *Gastrointest Endosc* 2021; 93: 183–186.
  34. Schroeder KW, Tremaine WJ and Ilstrup DM. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med* 1987; 317: 1625–1629.
  35. Travis SP, Schnell D, Krzeski P, et al. Reliability and initial validation of the ulcerative colitis endoscopic index of severity. *Gastroenterology* 2013; 145: 987–995.
  36. Rimondi A, Gottlieb K, Despott EJ, et al. Can artificial intelligence replace endoscopists when assessing mucosal healing in ulcerative colitis? A systematic review and diagnostic test accuracy meta-analysis. *Dig Liver Dis* 2024; 56: 1164–1172.
  37. Feagan BG, Rutgeerts P, Sands BE, et al. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013; 369: 699–710.
  38. Collaborators C-I, Ma C, Hanzel J, et al. CORE-IBD: A multidisciplinary international consensus initiative to develop a core outcome set for randomized controlled trials in inflammatory bowel disease. *Gastroenterology* 2022; 163: 950–964.
  39. Yao H, Najarian K, Gryak J, et al. Fully automated endoscopic disease activity assessment in ulcerative colitis. *Gastrointest Endosc* 2021; 93: 728–736 e1.
  40. Lo B, Liu Z, Bendtsen F, et al. High accuracy in classifying endoscopic severity in ulcerative colitis using convolutional neural network. *Am J Gastroenterol* 2022; 117: 1648–1654.
  41. Iacucci M, Cannatelli R, Parigi TL, et al. A virtual chromoendoscopy artificial intelligence system to detect endoscopic and histologic activity/remission and predict clinical outcomes in ulcerative colitis. *Endoscopy* 2022; 55: 332–341.
  42. Wang H, Jin Q, Li S, et al. A comprehensive survey on deep active learning in medical image analysis. *Med Image Anal* 2024; 95: 103201.
  43. Gottlieb K, Requa J, Karnes W, et al. Central reading of ulcerative colitis clinical trial videos using neural networks. *Gastroenterology* 2021; 160: 710–719.e2.
  44. Lobaton T, Bessissow T, De Hertogh G, et al. The Modified Mayo Endoscopic Score (MMES): a new index for the assessment of extension and severity of endoscopic activity in ulcerative colitis patients. *J Crohns Colitis* 2015; 9: 846–852.
  45. Stidham RW, Cai L, Cheng S, et al. Using computer vision to improve endoscopic disease quantification in therapeutic clinical trials of

- ulcerative colitis. *Gastroenterology* 2024; 166: 155–167 e2.
46. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) Initiative of the International Organization for the Study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. *Gastroenterology* 2021; 160: 1570–1583.
  47. Jairath V, Zou G, Wang Z, et al. Determining the optimal treatment target in patients with ulcerative colitis: rationale, design, protocol and interim analysis for the randomised controlled VERDICT trial. *BMJ Open Gastroenterol* 2024; 11: e001218.
  48. Ma C, Sedano R, Almradi A, et al. An International Consensus to standardize integration of histopathology in ulcerative colitis clinical trials. *Gastroenterology* 2021; 160: 2291–2302.
  49. Gui X, Bazarova A, Del Amor R, et al. PICaSSO Histologic Remission Index (PHRI) in ulcerative colitis: development of a novel simplified histological score for monitoring mucosal healing and predicting clinical outcomes and its applicability in an artificial intelligence system. *Gut* 2022; 71: 889–898.
  50. Najdawi F, Sucipto K, Mistry P, et al. Artificial intelligence enables quantitative assessment of ulcerative colitis histology. *Mod Pathol* 2023; 36: 100124.
  51. Rymarczyk D, Schultz W, Borowa A, et al. Deep learning models capture histological disease activity in Crohn's disease and ulcerative colitis with high fidelity. *J Crohns Colitis* 2024; 18: 604–614.
  52. Iacucci M, Parigi TL, Del Amor R, et al. Artificial intelligence enabled histological prediction of remission or activity and clinical outcomes in ulcerative colitis. *Gastroenterology* 2023; 164: 1180–1188 e2.
  53. Vande Casteele N, Leighton JA, Pasha SF, et al. Utilizing deep learning to analyze whole slide images of colonic biopsies for associations between eosinophil density and clinicopathologic features in active ulcerative colitis. *Inflamm Bowel Dis* 2022; 28: 539–546.
  54. Furlanello C, Bussola N, Merzi N, et al. The development of artificial intelligence in the histological diagnosis of Inflammatory Bowel Disease (IBD-AI). *Dig Liver Dis* 2025; 57: 184–189.
  55. Ohara J, Nemoto T, Maeda Y, et al. Deep learning-based automated quantification of goblet cell mucus using histological images as a predictor of clinical relapse of ulcerative colitis with endoscopic remission. *J Gastroenterol* 2022; 57: 962–970.
  56. Liu X, Prasath S, Siddiqui I, et al. Machine learning-based prediction of pediatric ulcerative colitis treatment response using diagnostic histopathology. *Gastroenterology* 2024; 166: 921–924 e4.
  57. Ohara J, Maeda Y, Ogata N, et al. Automated neutrophil quantification and histological score estimation in ulcerative colitis. *Clin Gastroenterol Hepatol* 2024; S1542-3565(24)00668-2.
  58. Peyrin-Biroulet L, Adsul S, Stancati A, et al. An artificial intelligence-driven scoring system to measure histological disease activity in ulcerative colitis. *United European Gastroenterol J* 2024; 12: 1028–1033.
  59. Iacucci M, Maeda Y and Ghosh S. Artificial intelligence enabled histological scoring in ulcerative colitis: are we ready yet? *United European Gastroenterol J* 2024; 12: 1000–1001.
  60. Bettenworth D, Baker ME, Fletcher JG, et al. A global consensus on the definitions, diagnosis and management of fibrostenosing small bowel Crohn's disease in clinical practice. *Nat Rev Gastroenterol Hepatol* 2024; 21(8): 572–584.
  61. Rieder F, Ma C, Hanzel J, et al. Reliability of CT enterography for describing fibrostenosing Crohn disease. *Radiology* 2024; 312: e233038.
  62. Rieder F, Baker ME, Bruining DH, et al. Reliability of MR enterography features for describing fibrostenosing Crohn disease. *Radiology* 2024; 312: e233039.
  63. 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: 734–742.
  64. 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: 8692–8705.
  65. Gu P, Chang JH, Carter D, et al. Radiomics-based analysis of intestinal ultrasound images for inflammatory bowel disease: a feasibility study. *Crohns Colitis* 360 2024; 6: otae034.
  66. Carter D, Albshesh A, Shimon C, et al. Automated detection of Crohn's disease in intestinal ultrasound using convolutional neural



- network. *Inflamm Bowel Dis* 2023; 29: 1901–1906.
67. Jongsma MME, Costes LMM, Tindemans I, et al. Serum immune profiling in paediatric Crohn's disease demonstrates stronger immune modulation with first-line infliximab than conventional therapy and pre-treatment profiles predict clinical response to both treatments. *J Crohns Colitis* 2023; 17: 1262–1277.
  68. Ananthakrishnan AN, Luo C, Yajnik V, et al. Gut microbiome function predicts response to anti-integrin biologic therapy in inflammatory bowel diseases. *Cell Host Microbe* 2017; 21: 603–610.e3.
  69. Sinonquel P, Bossuyt P, Sabino JPG, et al. Long-term follow-up of the red density pilot trial: a basis for long-term prediction of sustained clinical remission in ulcerative colitis? *Endosc Int Open* 2023; 11: E880–E884.
  70. 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: 1778–1786.
  71. Maeda Y, Kudo SE, Ogata N, et al. Evaluation in real-time use of artificial intelligence during colonoscopy to predict relapse of ulcerative colitis: a prospective study. *Gastrointest Endosc* 2022; 95: 747–756 e2.
  72. Hirten RP, Lin KC, Whang J, et al. Longitudinal monitoring of IL-6 and CRP in inflammatory bowel disease using IBD-AWARE. *Biosens Bioelectron X* 2024; 16: 100435.
  73. Hirten RP, Lin KC, Whang J, et al. Longitudinal assessment of sweat-based TNF-alpha in inflammatory bowel disease using a wearable device. *Sci Rep* 2024; 14: 2833.
  74. Ivanisevic T and Sewduth RN. Multi-omics integration for the design of novel therapies and the identification of novel biomarkers. *Proteomes* 2023; 11(4): 34.
  75. Christou CD and Tsoulfas G. Challenges involved in the application of artificial intelligence in gastroenterology: the race is on! *World J Gastroenterol* 2023; 29: 6168–6178.
  76. Hinton G. Deep learning—a technology with the potential to transform health care. *JAMA* 2018; 320: 1101–1102.
  77. Cohen-Mekelburg S, Berry S, Stidham RW, et al. Clinical applications of artificial intelligence and machine learning-based methods in inflammatory bowel disease. *J Gastroenterol Hepatol* 2021; 36: 279–285.
  78. Rajkomar A, Dean J and Kohane I. Machine learning in medicine. *N Engl J Med* 2019; 380: 1347–1358.
  79. Gubatan J, Levitte S, Patel A, et al. Artificial intelligence applications in inflammatory bowel disease: Emerging technologies and future directions. *World J Gastroenterol* 2021; 27: 1920–1935.
  80. D'Amico F, Danese S and Peyrin-Biroulet L. Adaptive designs: lessons for inflammatory bowel disease trials. *J Clin Med* 2020; 9: 2350.
  81. Ahmad HA, East JE, Panaccione R, et al. Artificial intelligence in inflammatory bowel disease: implications for clinical practice and future directions. *Intest Res* 2023; 21: 283–294.
  82. Dercle L, Fronheiser M, Lu L, et al. Identification of non-small cell lung cancer sensitive to systemic cancer therapies using radiomics. *Clin Cancer Res* 2020; 26: 2151–2162.
  83. Corbin CK, Sung L, Chattopadhyay A, et al. Personalized antibiograms for machine learning driven antibiotic selection. *Commun Med (Lond)* 2022; 2: 38.
  84. Kroner PT, Engels MM, Glicksberg BS, et al. Artificial intelligence in gastroenterology: a state-of-the-art review. *World J Gastroenterol* 2021; 27: 6794–6824.
  85. He M, Li C, Tang W, et al. Machine learning gene expression predicting model for ustekinumab response in patients with Crohn's disease. *Immun Inflamm Dis* 2021; 9: 1529–1540.
  86. Park SK, Kim YB, Kim S, et al. Development of a machine learning model to predict non-durable response to anti-TNF therapy in Crohn's disease using transcriptome imputed from genotypes. *J Pers Med* 2022; 12: 947.
  87. Waljee AK, Liu B, Sauder K, et al. Predicting corticosteroid-free biologic remission with vedolizumab in Crohn's disease. *Inflamm Bowel Dis* 2018; 24: 1185–1192.
  88. Con D, van Langenberg DR and Vasudevan A. Deep learning vs conventional learning algorithms for clinical prediction in Crohn's disease: a proof-of-concept study. *World J Gastroenterol* 2021; 27: 6476–6488.
  89. Qaiser T, Hamidinekoo A, Daniel S, et al. Artificial intelligence to predict interleukin-23 signalling activity from H&E stained whole slide images of inflammatory bowel disease. *J Crohn's Colitis* 2024; 18: i453.



90. Takenaka K, Fujii T, Kawamoto A, et al. Deep neural network for video colonoscopy of ulcerative colitis: a cross-sectional study. *Lancet Gastroenterol Hepatol* 2022; 7: 230–237.
91. Waljee AK, Wallace BI, Cohen-Mekelburg S, et al. Development and validation of machine learning models in prediction of remission in patients with moderate to severe Crohn disease. *JAMA Netw Open* 2019; 2: e193721.
92. Choi YI, Park SJ, Chung JW, et al. Development of machine learning model to predict the 5-year risk of starting biologic agents in patients with inflammatory bowel disease (IBD): K-CDM Network Study. *J Clin Med* 2020; 9: 3427.
93. Waljee AK, Lipson R, Wiitala WL, et al. Predicting hospitalization and outpatient corticosteroid use in inflammatory bowel disease patients using machine learning. *Inflamm Bowel Dis* 2017; 24: 45–53.
94. Wang Y, He Q, Yao D, et al. Histological image-based ensemble model to identify myenteric plexitis and predict endoscopic postoperative recurrence in Crohn's disease: a multicentre, retrospective study. *J Crohns Colitis* 2024; 18: 727 EP–737.
95. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 2019; 569: 655–662.
96. Lee JWJ, Plichta D, Hogstrom L, et al. Multi-omics reveal microbial determinants impacting responses to biologic therapies in inflammatory bowel disease. *Cell Host Microbe* 2021; 29: 1294–1304 e4.
97. Iacucci M, Santacroce G, Zammarchi I, et al. Artificial intelligence and endo-histo-omics: new dimensions of precision endoscopy and histology in inflammatory bowel disease. *Lancet Gastroenterol Hepatol* 2024; 9: 758–772.
98. Sahoo D, Swanson L, Sayed IM, et al. Artificial intelligence guided discovery of a barrier-protective therapy in inflammatory bowel disease. *Nat Commun* 2021; 12: 4246.
99. Maddox TM, Rumsfeld JS and Payne PRO. Questions for artificial intelligence in health care. *JAMA* 2019; 321: 31–32.
100. Zhou J, Hu N, Huang ZY, et al. Application of artificial intelligence in gastrointestinal disease: a narrative review. *Ann Transl Med* 2021; 9: 1188.
101. Prepared by AAITF, Parasa S, Berzin T, et al. Consensus statements on the current landscape of artificial intelligence applications in endoscopy, addressing roadblocks, and advancing artificial intelligence in gastroenterology. *Gastrointest Endosc* 2025; 101: 2–9.e1.
102. FDA. U.S. Food and Drug Administration. Evaluate application of artificial intelligence to adaptive enrichment clinical trials, <https://www.fda.gov/science-research/advancing-regulatory-science/evaluate-application-artificial-intelligence-adaptive-enrichment-clinical-trials> (2021, accessed 20 October 2024).
103. Zhang B, Zhang L, Chen Q, et al. Harnessing artificial intelligence to improve clinical trial design. *Commun Med (Lond)* 2023; 3: 191.
104. Zhu H and Wong WK. An overview of adaptive designs and some of their challenges, benefits, and innovative applications. *J Med Internet Res* 2023; 25: e44171.
105. Ravi D, Wong C, Deligianni F, et al. Deep learning for health informatics. *IEEE J Biomed Health Inform* 2017; 21: 4–21.
106. Shajari S, Kuruvnashetti K, Komeili A, et al. The emergence of AI-based wearable sensors for digital health technology: a review. *Sensors (Basel)* 2023; 23: 9498.
107. Chopra H, Annu, Shin DK, et al. Revolutionizing clinical trials: the role of AI in accelerating medical breakthroughs. *Int J Surg* 2023; 109: 4211–4220.
108. Vampana SB, Jayanthi ESS, Mary DA, et al. Artificial intelligence-driven patient monitoring for adverse event detection in clinical trials. *Int J Basic Clin Pharmacol* 2024; 13: 543–550.
109. Haug CJ and Drazen JM. Artificial intelligence and machine learning in Clinical Medicine, 2023. *N Engl J Med* 2023; 388: 1201–1208.
110. Yang G, Ye Q and Xia J. Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond. *Inf Fusion* 2022; 77: 29–52.
111. Ali S, Akhlaq F, Imran AS, et al. The enlightening role of explainable artificial intelligence in medical & healthcare domains: a systematic literature review. *Comput Biol Med* 2023; 166: 107555.
112. Amann J, Vetter D, Blomberg SN, et al. To explain or not to explain?—artificial intelligence explainability in clinical decision support systems. *PLoS Digit Health* 2022; 1: e0000016.
113. Abgrall G, Holder AL, Chelly Dagdia Z, et al. Should AI models be explainable to clinicians? *Crit Care* 2024; 28: 301.

114. Labkoff S, Oladimeji B, Kannry J, et al. Toward a responsible future: recommendations for AI-enabled clinical decision support. *J Am Med Inform Assoc* 2024; 31: 2730–2739.
115. Sadeghi Z, Alizadehsani R, Cifci MA, et al. A review of explainable artificial intelligence in healthcare. *Comput Electr Eng* 2024; 118: 109370.
116. Sedano R, Hogan M, McDonald C, et al. Underrepresentation of minorities and lack of race reporting in ulcerative colitis drug development clinical trials. *Inflamm Bowel Dis* 2022; 28: 1293–1295.
117. Sedano R, Hogan M, McDonald C, et al. Underrepresentation of minorities and underreporting of race and ethnicity in Crohn's disease clinical trials. *Gastroenterology* 2022; 162: 338–340.e2.
118. Shah S, Shillington AC, Kabagambe EK, et al. Racial and ethnic disparities in patients with inflammatory bowel disease: an online survey. *Inflamm Bowel Dis* 2024; 30: 1467–1474.
119. Nazer LH, Zatarah R, Waldrip S, et al. Bias in artificial intelligence algorithms and recommendations for mitigation. *PLoS Digit Health* 2023; 2: e0000278.
120. Arora A, Alderman JE, Palmer J, et al. The value of standards for health datasets in artificial intelligence-based applications. *Nat Med* 2023; 29: 2929–2938.
121. Rana SA, Azizul ZH and Awan AA. A step toward building a unified framework for managing AI bias. *Peer J Comput Sci* 2023; 9: e1630.
122. Yang J, Soltan AAS, Eyre DW, et al. Algorithmic fairness and bias mitigation for clinical machine learning with deep reinforcement learning. *Nat Mach Intell* 2023; 5: 884–894.
123. Mennella C, Maniscalco U, De Pietro G, et al. Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon* 2024;10:e26297.
124. Reddy S. Navigating the AI Revolution: the case for precise regulation in health care. *J Med Internet Res* 2023; 25: e49989.
125. Zhao Z, Alzubaidi L, Zhang J, et al. A comparison review of transfer learning and self-supervised learning: definitions, applications, advantages and limitations. *Expert Syst Appl* 2024; 242: 122807.
126. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ* 2023; 23: 689.
127. Hanley DF, Jr, Bernard GR, Wilkins CH, et al. Decentralized clinical trials in the trial innovation network: value, strategies, and lessons learned. *J Clin Transl Sci* 2023; 7: e170.
128. Daly B, Brawley OW, Gospodarowicz MK, et al. Remote monitoring and data collection for decentralized clinical trials. *JAMA Netw Open* 2024; 7: e246228.
129. Sharma A, Badea M, Tiwari S, et al. Wearable biosensors: an alternative and practical approach in healthcare and disease monitoring. *Molecules* 2021; 26.
130. Gelinis L, Morrell W, White SA, et al. Navigating the ethics of remote research data collection. *Clin Trials* 2021; 18: 606–614.
131. Gerke S, Minssen T and Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. In Bohr A and Memarzadeh K (eds) *Artificial intelligence in healthcare*, Denmark: Elsevier, 2020, pp.295–336.
132. Lu X, Chen M, Lu Z, et al. Artificial intelligence tools for optimising recruitment and retention in clinical trials: a scoping review protocol. *BMJ Open* 2024; 14: e080032.
133. Celi LA, Cellini J, Charpignon ML, et al. Sources of bias in artificial intelligence that perpetuate healthcare disparities—a global review. *PLoS Digit Health* 2022; 1: e0000022.
134. Noseworthy PA, Attia ZI, Brewer LC, et al. Assessing and mitigating bias in medical artificial intelligence: the effects of race and ethnicity on a deep learning model for ECG analysis. *Circ Arrhythm Electrophysiol* 2020; 13: e007988.
135. Aung YYM, Wong DCS and Ting DSW. The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *Br Med Bull* 2021; 139: 4–15.
136. WHO. WHO guidance. Ethics and governance of artificial intelligence for health, vol. 2024, <https://www.who.int/publications/item/9789240029200> (2021, accessed 20 October 2024).