Artificial Intelligence and Inflammatory Bowel Disease: Practicalities and Future Prospects

Authors
Johanne Brooks-Warburton*1,2; James J Ashton*3,4; Anjan Dhar5; Tony C Tham6, PB Allen6; Sami Hoque7 Laurence B. Lovat8, Shaji Sebastian9,10

*joint first author

Affiliations
1 Department of Clinical, Pharmaceutical and Biological Sciences, University of Hertfordshire, Hertford, UK
2 Gastroenterology Department, Lister hospital, Stevenage, UK
3 Department of Paediatric Gastroenterology, Southampton Children’s Hospital, Southampton, UK
4 Human Genetics and Genomic Medicine, University of Southampton, Southampton, UK
5 County Durham & Darlington NHS Foundation Trust and Teesside University
6 Department of Gastroenterology, Ulster Hospital, South Eastern Trust, Belfast, N. Ireland
7 Division of Surgery & Interventional Science, University College London, UK
8 Barts Health NHS Trust, London, UK
9 IBD Unit, Hull University Teaching Hospitals NHS Trust, Hull, UK
10 Hull York Medical School, Hull, UK.
Contributorship Statement

JBW organised the authors, did the literature review wrote, edited and revised the manuscript.

JJA did the literature review, wrote and revised the manuscript and designed all the figures.

AD wrote the manuscript, edited and revised the manuscript.

TCT did the literature review and reviewed the manuscript for content and edited.

PBA reviewed the manuscript for content and edited.

SH reviewed the manuscript for content and edited.

LL wrote the manuscript and provided significant editing input.

SS wrote and revised the manuscript and provided significant editing input.

Keywords
Inflammatory Bowel Disease, Artificial Intelligence, Genomics, Microbiome, Endoscopy, Apps, Internet of Things, Ethics
COIs

**JBW** holds research grants from AbbVie, and has received speaker fees from Dr Falk Pharma.

**JJA** no COI to declare

**AD** Advisory and consultancy for Pfizer UK, Tillotts Pharma UK, Dr Falk Pharma, Pharmacosmos, Takeda UK

Honoraria and Speaker Fees from Dr Falk Pharma, Pfizer, Janssen, Tillotts Pharma, Pharmacosmos, Takeda UK

**TCT** Honoraria from Galapagos, Amgen, Takeda, Janssen.

**PBA** received research grants from Pfizer and speaker fees from Takaeda, AbbVie, GSK, Janssen, MSD and Tillotts.

**LL** is a minor shareholder in Odin Vision since 2018

**SS** holds research grants from Biogen, Takeda, AbbVie, Tillotts Pharma, Ferring and Biohit; served on the advisory boards of Takeda, AbbVie, Merck, Ferring, Pharmacosmos, Warner Chilcott, Janssen, Falk Pharma, Biohit, TriGenix, Celgene and Tillotts Pharma; and has received speaker fees from AbbVie, Biogen, AbbVie, Janssen, Merck, Warner Chilcott and Falk Pharma

**Key Points:**

- AI is a growing field which could improve patient care in IBD
- AI in IBD uses two main types: Machine learning and Deep Learning
- AI is currently used for IBD pathogenesis, drug discovery and endoscopic identification of disease
- AI could be used in IBD patient self-monitoring, virtual management, endoscopic disease classification and patient stratification
- Narrow, task-focused artificial intelligence methodologies can and do have a positive impact on patient care
Abstract

Artificial intelligence is an emerging technology predicted to have significant applications in health care. This review highlights AI applications that impact the patient journey in Inflammatory Bowel Disease, from genomics to endoscopic applications in disease classification, stratification and self-monitoring to risk stratification for personalized management. We discuss the practical artificial intelligence applications currently in use whilst giving a balanced view of concerns and pitfalls and look to the future with the potential of where AI can provide significant value to the care of the IBD patient.

Introduction

Artificial intelligence (AI) is a computational system designed by humans that, given a specific goal, takes in data, interprets it, processes it and decides on the best action to achieve the goal.

General AI systems perform non-biological tasks that humans can do. To do this, they require the capabilities of common sense, self-awareness or the ability to define their purpose. They are not widely used in medicine. Narrow AI systems are those that we recognise in everyday use, which has one function, e.g. AI optimising hospital bed flow.

AI encompasses several approaches, the commonest being Machine Learning (ML). This applies computer algorithms to capture behaviour and patterns in systems and processes based on the input and output data. An example of this is natural language processing (NLP), in which a computer is trained to understand text and spoken word. Within ML, there is a specific subset called deep learning (DL). DL utilises neural networks with multiple layers, that is, networks that aim to mimic the brain through a combination of extensive data inputs, weights and biases (Figure 1). An example of this is Convolutional Neural Networks (CNNs). They are used widely in image-based problems in medicine, e.g. to detect and classify colonic polyps, to diagnose retinal disorders and to predict severity in chest x-rays for patients with congestive cardiac failure. Other approaches in AI DL which are less commonly used in medicine are machine reasoning and robotics. These are used in self-driving cars, but the idea of a machine diagnosing a patient with acute fulminant colitis, deciding to operate and performing a colectomy without any human intervention remains science fiction.
Inflammatory bowel disease is a chronic debilitating condition, which has significant uncertainties along the patient journey from pre-diagnosis to the maintenance of disease control. In this review, we highlight the current and potential future applications of AI in the IBD journey (Figure 2).

**Understanding IBD Pathogenesis**

A significant role for AI in IBD is to aid understanding of aetiology and personalise disease management, via deeper insights into the complexity of this chronic immune condition which has both heritable and environmental components (1).

**Genomics and Multiomics**

Interpretation of vast numbers of genetic variants in IBD and correlating these with phenotypes is highly challenging. This already utilises elements of AI; in the field of genetic variant annotation, a notable example is the Combined Annotation-Dependent Depletion (CADD) score (2). This method uses a ML model to assign a score to any genetic variant based on the determination of whether the variant will be deleterious or not. Other examples of AI use for variant annotation occur in the paediatric field for very early onset IBD patients (aged <6 years at diagnosis) who routinely undergo genomic sequencing (3). ML models such as LEAP utilise big data to identify paediatric IBD patients' deleterious genetic markers with excellent accuracy (AUC >0.97), providing a direct clinical impact of AI in IBD (4).

**The microbiome**

The microbiome is increasingly recognised to play a role in IBD and is very well suited to AI analyses. 'Big data metagenomics' utilising both ML and specific DL algorithms can distinguish the microbiomes of IBD and control populations (5). The use of AI within host-microbiome metabolomic networks affords a predictive insight into the microbiome-host interplay in the gut covering the spectrum of synergistic effects to dysbiosis in health and disease states (6).
A challenge of data integration lies in the heterogeneity of ‘omic data inputs, from genotypes, transcript counts, bacterial species abundance, and clinical metadata. This complexity is exceptionally hard to understand, but AI systems appear successful e.g., Hyams et al. used a combination of microbial, transcriptomics and clinical data to predict 52-week steroid-free remission in paediatric UC (7). Here the authors utilised a multivariable logistic regression model and identified an Area Under the Curve of 0.75 when using disease activity, week 4 remission, antimicrobial transcriptomic signature and specific bacterial abundance as input variables.

**AI in IBD Diagnosis:**

**Risk Prediction, Disease Classification**

Most of the existing literature on the use of AI in diagnosing IBD relates to developing risk prediction models using either ML or CNNs to analyse endoscopic and imaging datasets (8). But ML algorithms can also successfully classify disease subtypes based on genotype. Wei et al. utilised ML to classify CD and UC, vs controls (AUC = 0.862 and 0.826, respectively) (9) based on single nucleotide polymorphisms (SNPs). The Crohn's disease exome challenge applied ML and DL to SNP and whole-exome sequencing to produce a classification model with an AUC of 0.72 (10). Smolander et al. have explored the use of deep belief networks (DBNs), which are another type of neural network, in distinguishing ulcerative colitis from Crohn's disease with an accuracy of around 97% (11).

**AI for Endoscopic Assessment**

An area where most endoscopists value the potential for AI is endoscopic severity assessment. A recent prospective study using CNN in 875 patients with UC demonstrated the ability to detect endoscopic remission (Mayo Endoscopic Score, MES score of 0 vs 1) with 90% accuracy and histological remission with 92.9% accuracy comparable to expert reviewers (12). In a further study, AI was able to detect MES Scores
from both still and video images (13). Similar data is available for computer-aided diagnosis of grade 1 vs grade 3 ulcers in CD and histological inflammation by endocytoscopy. Takenaka et al. have designed a deep neural network model for UC (DNUC), trained on >40,000 endoscopic still digital images with corresponding ulcerative colitis endoscopic index of severity (UCEIS) scores and histology. Compared with human reviewers, this system was highly accurate in assessing endoscopic and histologic remission.

A recent proof of concept study using an automated real-time operator-independent red density score in ulcerative colitis reported a good correlation between the red density score, a consensus Mayo Endoscopic Score and UCEIS, as well as a Robart's histological index. (14). As a note of caution, translating accuracy from still images to live video analysis is a work in progress but solutions are starting to appear.

Other potential areas for application are standardised reporting of IBD endoscopy such as auto report generation through computer-aided detection (CAD) or task automation such as data extraction from electronic medical records, generating reports and billing through NLP.

Clinical trials in IBD increasingly use central reading to assess endoscopic severity. Automated endoscopic scoring in high-resolution trial endoscopy videos with frame integration using recurrent neural network showed high accuracy in prediction of the correct Mayo score (70%) with a satisfactory agreement with human observers (k) of 0.84 (0.79-0.90) (15). However, Yao et al. (13) showed that a high-quality video source is essential for optimal AI performance to be applicable in the setting of central reading. AI can also provide standardised and reproducible disease assessments and integrate images into decision support in strategy trials challenging the conventional therapeutic endpoints.

Capsule endoscopy images also benefit from AI to varying degrees. Barash et al. developed CNN to evaluate ulceration in CD in which the overall agreement between expert consensus and an automated algorithm was 67%(16). A DL model in capsule
endoscopy reading was superior to conventional reading both at the level of per-patient sensitivity (99.9% vs 74.6%, p=<0.0001) and per lesions sensitivity 99.9% vs 76.9%, p=<0.0001) (17).

EndoBRAIN-EYE (Cybernet Systems, Tokyo, Japan) has developed a prototype AI-based polyp detection system to detect flat dysplasia in long-standing colitis patients.

Al in Radiological Diagnosis

Al can also interpret IBD radiological images. Stidham et al. reported that structural bowel damage measurements of computed tomography enterography (CTE) data in CD by semi-automated approaches are comparable to those of experienced radiologists (18). Furthermore, radiomics models can predict moderate to severe histological intestinal fibrosis more accurately than skilled radiologists (19).

Therapy and Patient-focused management:

Al and Apps

Al-based healthcare apps are a growing industry, with context-aware smartphone applications for personalised outcomes for the user. Most healthcare apps use a mobile expert system model: a knowledge base containing the facts and rules relevant to the target application, which can be accessed and updated by an IBD Specialist, and a user interface, into which users place queries or results, e.g. stool frequency, level of pain, mood etc. The knowledge base and user interface are input into an inference engine that applies an IF-THEN rule (IF being the condition and THEN being the consequence), which feeds back to the user and the expert. An example of this in IBD is when a user places stool frequency at > 5 x normal, with blood - the app would advise doing a faecal calprotectin and contacting the IBD service. Some apps feedback to the expert interface to alert them of a change in circumstance.
A scoping review of Digital Health apps for the Clinical care of IBD (20) identified ten digital health apps for IBD, which utilised the app as a self-reporting/documentation app and provided educational modules. Half of the apps had a contact through to team capability, which was user instigated, with thresholds set for investigation, e.g. faecal calprotectin. One utilised ML algorithms to predict the requirement for escalation (with 95% accuracy), with treatment advice given regarding 5ASA and rectal therapy.

Like other health apps, IBD apps have the potential to provide patient-centred and patient-initiated care, improve access to specialised medical services in areas that are underserved in a low-cost way. However, complexities arise with the need for encryption for patient identifiable data and the discriminant function of symptoms in identifying non-IBD pathology such as cancer in IBD patients.

**AI in IBD Monitoring:**

**Chatbots**

Conversational AI methods such as chatbots may enhance timely access to IBD specialists. In using a chatbot, the input (query) must be able to be categorised effectively. In a feasibility study, an NLP algorithm was developed to classify IBD patient electronic messaging dialogue into specific queries: symptoms based, medication queries, investigation results, insurance, procedures, communications and miscellaneous. The algorithm had 90% concordance with two Consultant Gastroenterologists who reviewed the data and categorised it (21).

**Smart Loos**

With the advent of the Internet of Things (IoT), e.g. the embedding of computing devices into everyday objects that connect via the internet, the opportunity to utilise everyday objects as a healthcare adjunct is arising. For IBD sufferers, the potential to use the toilet as a monitoring system for their disease is a valid and appropriate use of a combination of IoT and AI. Smart Toilets have been around since 2000, when they were initially capable of monitoring the glucose level in urine. The level of sophistication has expanded,
and Smart toilets can now categorise, by DL, the stool type (Bristol Stool scale) and longitudinally log frequency and type for the user (identified by their fingerprint and anoderm). This data is accurate biometric data for the physician (22). Fisher et al. (23) adapted the Smart Toilet to integrate an AI-based stool analysis (using CNN for stool type) and a perceptual colour quantisation for accurate real-time stool and faecal blood analysis. This has reduced recall inaccuracies and the burden of manually checking every bowel movement.

Wearables and Nearables

Worn monitoring devices called wearables and nearables (neighbouring devices that interact with wearables such as smartphones) are now mainstream. These are the most likely technologies to transform future healthcare and lifestyles via self-motivation management (24). Sensors to detect heart rate and step count are commonplace with integrated AI to promote healthy exercise. The value of exercise in managing fatigue and reducing the risk of flares in IBD is clear (25), and in this way, AI and IoT enabled wearables and nearables will be part of the IBD journey. The prospect of wearing a sensor for early detection of a flare via sweat is an exciting wearable development for IBD, but as yet had not been AI-enabled or IoT integrated (26).

AI in Drug Design

AI is used in all drug design and discovery facets. Traditionally Quantitative Structure-activity relationships (QSAR) were used, which made accurate predictions of simple attributes such as solubility, which are straightforward. More complex attributes such as drug efficacy or side effect identification were not plausible using this technique.

Using Big Data and large databases, AI modelling algorithms can predict interactions between drugs and biological targets. Wen et al. (27) used DL to identify drug-target interactions from DrugTarget Database. Xie et al. (28) utilised DL-based transcriptome data classification for drug-target interaction prediction from the Library of Integrated Network-Based Cellular Signatures to identify novel targets for known drugs.
Unbiased AI-assisted methods for target identification in IBD (29) identified relevant gene clusters in IBD and actionable targets. The AI algorithm identified PRKAB1 as a barrier protective therapeutic target, which was then experimentally validated. The model was able to predict phase III results in clinical trials in IBD successfully.

**AI to Predict Drug Response**

It is becoming possible to predict response to drugs, which is a required step for personalised medicine (30,31), e.g, Morilla et al using deep neural networks found nine colonic microRNAs with five clinical features that predicted response to steroids (AUC 0.91) and infliximab (AUC 0.82) (32). However, the use of integrated AI within these analyses is still in its infancy. It is essential for accurate clinical data to be used in these models and to validate findings in independent cohorts, as previous molecular biomarkers have not produced consistently reproducible results in all populations (33).

**AI Risks**

The difficulty with DL modelling is whilst it allows modelling of biological endpoints, which reduces the overfitting, biased data and errors from small datasets, there are no universally agreed tunable parameters or criteria for the workflows, which opens the models to criticism (34).

A recent safety issue highlighted by IBM Watson Oncology AI algorithm, which utilised synthetic data to train the algorithm, led to incorrect and unsafe therapeutic advice, albeit not to an actual patient. However, this situation has cast a shadow over medical apps providing feedback to patients regarding therapeutic strategies.

CNN's are used to analyse images produced from high throughput screening and are also used to visualise 3D experimental and virtual images for ligand-protein prediction (35). However, as such a wide application, it is fraught with difficulty regarding the available knowledge regarding the chemical space and the complexity of the biological context.
There are at least 10 AI algorithms/ML or DL applications for structural and ligand-based virtual screening, including LS-Align, LigGrep, DrugFinder and DEEPScreen (36).

Despite the immense potential, the drug discovery field lags behind other successes utilising AI in terms of the amount of chemical and biological data available, the ability to represent the data, the presence of a suitable stable algorithm, or the ability to appropriately assign biological labels, e.g. dose, genotype, age etc. (37).

**AI and Risks to the Doctor-Patient Relationship**

Arguably, the most important aspect of the doctor-patient relationship is trust. But AI lacks transparency because of the 'black box' nature of some algorithms. This leads to concerns about indefensible decisions and the loss of trust of the medical profession, potentially reducing patient engagement and concordance with therapy. Demystifying AI and clarity about where AI can aid the physician instead of replacing the physician will help allay such fears. Box 1 identifies current ethical and safety aspects of AI.

**Training**

In a recent UK based qualitative study, perception of AI in the medical community was worryingly one of fear and concern about ethical consequences, despite two-thirds of the participants never having used AI applications. Nearly 90% had no working knowledge of the principles of AI (38). Funding for education and engagement is going to be crucial so that clinicians understand how to use AI appropriately and appreciate its limitations. All medical AI applications must be GDPR compliant, with clear accountability and confirmation of privacy protection.

Perhaps even more important than training clinicians is to train the AI itself. Quality assurance depends on the quality of the training data sets. An AI that was not trained on gastric cancer is highly unlikely to recognise it and there may be subtle differences between patients that we are not aware of including age, sex and race differences. Transparent, standardised thresholds need to be set. The narrow nature of AI is both a
blessing and a curse. e.g., an algorithm designed to characterise IBD subtypes will not aid a trainee in loop resolution. But imagine having an AI that can suggest interventions to the junior endoscopist that lead to safer colonoscopy when loops form.

Conclusions:

This review has highlighted areas of the IBD journey that already benefit from AI in a tangible, practical sense and suggested other areas where AI is plausible but prospective. In doing so, our aim is to familiarise, raise awareness, dispel concerns, and enthuse medical colleagues as to the utility of AI within IBD. The success of AI integration in the IBD journey will be measured by the value-added effect; increased understanding of the disease, better diagnostic certainty, efficacy improvements, better holistic outcomes for the patients, and improvements in the work-life balance of the IBD team. AI systems are being developed to give practitioners extra resources to provide better care.

The concept of the broad AI unit, which thinks, acts, responds and is self-aware, is rife with ethical and moral issues and is beset by the vilification ingrained within the literary and movie genres. Replacing the medical practitioner with an AI module in everyday practice remains a slightly distant concept. This review shows where narrow, task-focused AI methodologies can and do have a positive impact on patient care. Within IBD; Genomics Research, Endoscopy image classification and to a lesser extent, patient-facing home monitoring apps have led the way to confirm the place of AI in the IBD patient journey.

Overall we predict the role of AI will continue to grow in IBD and may help to address some of the clinical challenges we face, leading to improved quality in IBD care.
Glossary of terms

**Algorithm:** The process or rules to be followed in problem-solving operations

**Blackbox:** A complex system whose internal workings are hidden or not easily understood

**Chatbot:** A computer program designed to simulate conversation, usually over the internet.

**Convoluted neural network (CNN):** Multi-layered networks, specialising in image processing, designed to detect multiple features in images to correctly identify the image - e.g. social media ‘automatic tagging’ of individuals in photos

**Deep Learning:** Class of machine learning that uses artificial neural networks (such as CNNs) to extract progressively high-level data from raw data e.g. identification of a type of polyp in an endoscopic image

**DeepScreen:** Large scale drug-target interaction prediction system using deep CNNs (can be found on github.com)

**DrugFinder:** Drug Target Interaction (DTI) database

**Inference Engine:** Processing component which makes a decision from facts/rules contained in the knowledge base of an expert system.

**Internet of Things (IoT):** System of interrelated, internet connected objects e.g. a fridge which orders the weekly shopping and links to your smartphone

**LEAP:** LEArn to Prescribe algorithm is a sequential decision making process to identify adverse drug reactions in the state of polypharmacy

**LS-Align:** A tool for screening large numbers of compounds against disease relevant proteins

**Machine learning:** The use and development of computer systems that are able to learn and adapt by using algorithms and models to analyse and draw inferences from patterns in data.

**Natural language processing:** Field of linguistics, computer science and AI to train a computer to understanding language, including nuances, to be able to extract large amounts of information from documents.

**Neural Networks:** Series of algorithms that recognize underlying relationships in a set of data through processes that mimic the connections of neurons in the brain.
**QSAR** (Quantitative Structure Activity Relationship): Theoretical models that relate a quantitative measure of a chemical structure to a physical property.

**Useful sites to find out more:**

- **AI innovation for NHS specific funding:**
  - NHSx [The Artificial Intelligence in Health and Care Award - NHSX](https://www.nhsx.nhs.uk/)

- **Natural Language processing:**
  - § Free resources: MIT, Edureka, DataScienceDojo, FreeCodeCamp
    - YouTube search term Natural Language Processing
  - § [Learn How to Get Started with Natural Language Processing | Codecademy](https://www.codecademy.com/customers/)

- **Deep Learning:**
  - Deep Learning | Coursera
  - Deep Learning Courses - Master Neural Networks, Machine Learning, Data Science, and Artificial Intelligence in Python, TensorFlow, PyTorch, and Numpy

- **Machine Learning**
  - www.datacamp.com
  - Machine Learning Crash Course | Google Developers
  - Machine Learning with Python: A Practical Introduction | edX

- **R:**
  - R: The R Project for Statistical Computing (r-project.org)
  - RStudio | Open source & professional software for data science teams - RStudio
  - R Courses & Tutorials | Codecademy
  - Training - RStudio

- **Python**
  - Learn Python - Free Interactive Python Tutorial
Bibliography


