



Universiteit
Leiden
The Netherlands

Combining mechanistic modeling with machine learning as a strategy to predict inflammatory bowel disease clinical scores

Shim, J.V.; Rehberg, M.; Wagenhuber, B.; Graaf, P.H. van der; Chung, D.W.

Citation

Shim, J. V., Rehberg, M., Wagenhuber, B., Graaf, P. H. van der, & Chung, D. W. (2025). Combining mechanistic modeling with machine learning as a strategy to predict inflammatory bowel disease clinical scores. *Frontiers In Pharmacology*, 16. doi:10.3389/fphar.2025.1479666

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/4249662>

Note: To cite this publication please use the final published version (if applicable).



OPEN ACCESS

EDITED BY

H. Christian Weber,
Boston University, United States

REVIEWED BY

Mathieu Garand,
Washington University in St. Louis, United States
Tiyash Parira,
Florida International University, United States

*CORRESPONDENCE

Douglas W. Chung,
✉ doug.chung@certara.com

RECEIVED 12 August 2024

ACCEPTED 27 January 2025

PUBLISHED 25 February 2025

CITATION

Shim JV, Rehberg M, Wagenhuber B,
van der Graaf PH and Chung DW (2025)
Combining mechanistic modeling with
machine learning as a strategy to predict
inflammatory bowel disease clinical scores.
Front. Pharmacol. 16:1479666.
doi: 10.3389/fphar.2025.1479666

COPYRIGHT

© 2025 Shim, Rehberg, Wagenhuber, van der
Graaf and Chung. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Combining mechanistic modeling with machine learning as a strategy to predict inflammatory bowel disease clinical scores

Jaehee V. Shim¹, Markus Rehberg², Britta Wagenhuber²,
Piet H. van der Graaf^{1,3} and Douglas W. Chung^{1*}

¹Certara Applied BioSimulation, Sheffield, United Kingdom, ²Sanofi R&D, Translational Disease Modeling, Frankfurt amMain, Germany, ³Division of Systems Pharmacology and Pharmacy, Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

Disease activity scores are efficacy endpoints in clinical trials of inflammatory bowel disease (IBD) therapies. Crohn's disease activity index (CDAI), Mayo endoscopic score (MES) and Mayo score are frequently used in clinical trials. They rely on either the physician's observation of the inflammatory state of the patient's gastrointestinal tissue alone or combined with the patient's subjective evaluation of general wellbeing. Given the importance of these scores in evaluating the efficacy of drug treatment and disease severity, there has been interest in developing a computational approach to reliably predict these scores. A promising approach is using mechanistic models such as quantitative systems pharmacology (QSP) which simulate the mechanisms of the disease and its modulation by the drug pharmacology. However, extending QSP model simulations to clinical score predictions has been challenging due to the limited availability of gut biopsy measurements and the subjective nature of some of the evaluation criteria for these scores that cannot be described using mechanistic relationships. In this perspective, we examine details of IBD disease activity scores and current progress in building predictive models for these scores (such as biomarkers for disease activity). Then, we propose a method to leverage simulated markers of inflammation from a QSP model to predict IBD clinical scores using a machine learning algorithm. We will demonstrate how this combined approach can be used to (1) explore mechanistic insights underlying clinical observations; and (2) simulate novel therapeutic strategies that could potentially improve clinical outcomes.

KEYWORDS

inflammatory bowel disease (IBD), machine learning, mayo score prediction, endoscopic score prediction, quantitative systems pharmacology (QSP) model, CDAI score prediction

Introduction

Inflammatory bowel disease (IBD) is a chronic inflammatory gut disease prevalent in the United States with approximately 1.6 million residents affected and over two million are estimated to be suffering in Europe (Gajendran et al., 2018; Ramos and Papadakis, 2019). IBD is classified into two subtypes (1) Crohn's disease (CD) which can affect anywhere from the mouth to the perianal area and (2) ulcerative colitis (UC) primarily affects the large intestine (Gajendran et al., 2018; Gajendran et al., 2019; Rogers et al., 2021a). Both types of

IBD can present with complications such as fistulae and strictures which may require surgery (Kim et al., 2013; Kishi et al., 2022; Rogers et al., 2021a; Venkatapurapu et al., 2022).

The pathogenesis of IBD involves environmental factors that can influence the changes in the microbiome and genetic factors that increase the susceptibility to gut inflammation. The disease mechanism involves a pathogenic microbiome entering through a weakened intestinal barrier which then leads to a dysregulated mucosal immune response ending in relapsing-remitting gut inflammation (Ahluwalia et al., 2018; Ramos and Papadakis, 2019; Sartor, 2006).

Many approved IBD therapeutics target immunological dysregulation to reduce gut inflammation and ameliorate symptoms. Immunomodulatory drugs that have been shown to clinically improve IBD symptoms include (1) tumor necrosis factor (TNF) antibodies (infliximab, adalimumab, golimumab), (2) interleukin (IL) 12/23 antibodies (ustekinumab, mirikizumab, guselkumab, risankizumab, brazikumab), (3) integrin antibodies (carotegrast, vedolizumab), and (4) Janus kinase (JAK) inhibitors (tofacitinib, filgotinib, upadacitinib) (Kobayashi and Hibi, 2023; Rogers et al., 2021a).

To better understand the disease biology and optimize therapeutic strategies, there has been significant interest in developing quantitative systems pharmacology (QSP) models that can capture the interactions between different immune cells, cytokines, interleukins, epithelial barrier, and gut microbiome (Balbas-Martinez et al., 2018; Fendt et al., 2024; Lo et al., 2013; Pinton, 2022; Pinton, 2023; Rogers et al., 2021b; Rogers et al., 2021a; Stübler et al., 2023; Venkatapurapu et al., 2022; Wendelsdorf et al., 2010; Whittaker et al., 2024). Because model connections are calibrated to available data, typically the model output is in the form of clinically measurable biomarkers such as fecal calprotectin (FCP) and serum c-reactive protein (CRP) (Rogers et al., 2021a; Rogers et al., 2021b; Venkatapurapu et al., 2022). However, in clinical trials, clinical scores are preferred metrics over biomarkers such as FCP and serum CRP because, while biomarkers are useful indicators to infer endoscopically active IBD, their levels do not necessarily correlate with disease activity level (Falvey et al., 2015; Wagatsuma et al., 2021). To reliably predict IBD clinical scores, there needs to be sufficient training data relating gut inflammation to measurable clinical markers and clinical scores. The IBD QSP model can generate simulated data on gut immunocytes and cytokine levels overcoming the limited quantity of patient data. The simulated patient data can be utilized to train a machine-learning model to describe the relationships between gut inflammatory markers and IBD clinical scores.

Here, we present a computational strategy to reliably generate comprehensive clinical score predictions for both UC and CD using gut-level biomarker simulations with a published IBD QSP model (Rogers et al., 2021a; Rogers et al., 2021b). We will showcase how such an integrated model can be used to explore (1) potential mechanistic differences behind responders versus non-responders of anti-TNF α and anti-IL-23 combination therapy in UC patients (inspired by a recent study by Feagan et al. (2023)); and (2) if the same dual combination would be beneficial for the other IBD subtype, CD.

IBD clinical scores

Clinical scores are utilized in clinical practice and in drug development to assess disease severity and can be categorized into (1) overall activity score which is a comprehensive measurement of disease severity, and (2) endoscopic score which is solely based on physical symptoms observed during endoscopy. For each subtype of IBD, there are various standardized scores for both overall and endoscopic assessment.

For UC, the most popular index for overall disease activity is the Mayo score (partial Mayo included). For clinical trials conducted from 2013 to 2017, about half of the trials (49.5%) utilized the Mayo score as clinical index to evaluate therapeutic responses (Kishi et al., 2022). Related indices include the Clinical Activity Index (CAI), Disease Activity Index (DAI), and Simple Clinical Colitis Activity Index (SCCAI). The most popular choice for the endoscopic index for UC was the Mayo Endoscopic Score (MES), which was employed in 69.0% of the clinical trials conducted from 2013 to 2017 (Kishi et al., 2022). Additional commonly used endoscopic indices are the Sutherland Endoscopic Sub-score and Ulcerative Colitis Endoscopic Index of Severity (UCEIS).

In CD, the most used disease activity index in clinical trials for overall assessment was the Crohn's Disease Activity Index (CDAI). Alternatively, the Harvey-Bradshaw Index (simple CDAI) is also utilized. For the endoscopic evaluation of CD, the CD Endoscopic Index of Severity (CDEIS) has been noted as the gold standard. Other options that have been utilized in place of the CDEIS are the Simple Endoscopic Score for Crohn's disease (SES-CD) and the Rutgeert Score (Kim, 2022; Kishi et al., 2022).

The most popular clinical indices in each category for CD (CDAI for overall, CDEIS for endoscopic score) and UC (Mayo for overall, MES for endoscopic score) are reviewed in more detail.

CDAI

The CDAI was first introduced in 1976 as a part of the National Cooperative Crohn's Disease Study and has since been considered the gold standard for evaluating CD (Kishi et al., 2022; Liu and Lichtenstein, 2012). To assign a CDAI score, there are eight items evaluated over 7 days (Freeman, 2008; Kishi et al., 2022; Liu and Lichtenstein, 2012). Some items are highly subjective such as abdominal pain score in 1 week and general wellbeing. Other factors considered are the number of liquid or very soft stools, the sum of physical findings per week (fever, anal disease, mucocutaneous lesions, arthritis, external fistula), antidiarrheal use, abdominal mass, and low hematocrit count. Each of the evaluation criteria is multiplied by a weighting factor and summed to derive the score (Freeman, 2008; Kishi et al., 2022). The CDAI scores below 150 are interpreted as no disease activity, 150–220 as mild, 220–450 as moderate, and >450 as severe (Nakamura et al., 2018; Shinzaki et al., 2021). For the clinical trials conducted from 2009 to 2017, the CDAI was reported to be the most frequently used index in over 50% of trials (Kishi et al., 2022).

CDEIS

The CDEIS was developed in 1989 by the Groupe d'Etudes Therapeutiques des affections Inflammatoires du tube Digestif. Items considered for the CDEIS scoring include the presence of superficial or deep ulcerations and the percentage of ulcerated surface. Scores can range from 0 to 44 and can be interpreted as healed if the score is 0–3, mild disease 4, moderate disease 5–15, and severe disease > 15 (Kim, 2022). The CDEIS has been the most popular choice in clinical trials for 2009–2012 while the SES-CD has been preferred in more recent trials (Kim, 2022; Kishi et al., 2022).

Mayo

The Mayo score was published in 1987 by Schroeder et al. in a study to evaluate the therapeutic effects of coated oral 5-aminosalicylic acid for UC (Kishi et al., 2022; Schroeder et al., 1987). Since then, the Mayo score has been the most utilized index in clinical trials, but its validity has not been examined thoroughly (Kishi et al., 2022). Scoring is based on four items which include stool frequency, rectal bleeding, mucosal appearance at endoscopy and physician rating of disease activity. Each item is rated between 0 and 3 and the final score is derived by adding up the ratings from each category (J. D. Lewis et al., 2008). The Mayo score is interpreted as remission if the score is 0–2, mild if 3–5, moderately active if 6–10 and severely active if 11–12 (CADTH, 2016).

MES

The MES was developed as a component for the Mayo score by Schroeder et al. in the 1987 study (Schroeder et al., 1987). It is evaluated in four points, ranging from 0–3, where the MES score of 0 represents normal and 3 represents an ulcerated state (Kim, 2022; Kishi et al., 2022). The descriptors utilized in the MES assessment include erythema, vascularity, friability, bleeding, erosions, and ulcerations (Kim, 2022). While the MES is the most popular endoscopic index for recent clinical trials, its validity has not been sufficiently evaluated (Kim, 2022; Kishi et al., 2022).

Challenges of and progress toward extending the QSP model to predict IBD clinical scores

Mechanistic modeling of IBD has been valuable in understanding the disease pathophysiology, exploring therapeutic targets, and optimizing treatment strategies (Balbas-Martinez et al., 2018; Kilian et al., 2024; Lo et al., 2013; Pinton, 2022; Pinton, 2023; Rogers et al., 2021a; Stübler et al., 2023; Venkatapurapu et al., 2022; Wendelsdorf et al., 2010; Yu et al., 2024). Published IBD models include (1) microbiome and epithelial barrier dysfunction, (2) immunological dysregulation in lamina propria and lymph nodes involving innate immune cells (neutrophils, macrophages, dendritic cells, natural killer T cells) and adaptive immune cells (B cells and T cells) (Abraham and Cho, 2009; Cai et al., 2021; Kilian et al., 2024; Saez et al., 2023; Stübler et al., 2023; Yu et al., 2024). Because many of

the approved drugs target distinct immunological pathways, there is increasing interest in finding synergistic combinations of existing drugs to optimize therapeutic outcomes (Dai et al., 2023; Feagan et al., 2023; Pinton, 2022; Wetwittayakhleng and Lakatos, 2024). For instance, a recent study by Feagan et al. showed that dual targeting of TNF α and IL-23 almost doubled the clinical remission rate compared to single-targeting therapies (Feagan et al., 2023). Other effective combinations include ustekinumab and vedolizumab (Dawoud et al., 2022), infliximab and azathioprine (Sultan et al., 2017), as well as adalimumab and vedolizumab (Goessens et al., 2021).

Identifying optimal dual targeting strategies has become the main application of QSP IBD modeling. However, many of the published models (Abraham and Cho, 2009; Cai et al., 2021; Kilian et al., 2024; Saez et al., 2023; Stübler et al., 2023; Yu et al., 2024) lack the extension to predict clinical scores. Having the capacity to simultaneously simulate changes in gut-level dynamics of cytokines and immune cells along with clinical scores enables a mechanistic understanding of treatment response by closely matching virtual populations with the biomarker and endpoint response in real-life clinical trials.

Although there is great interest in predicting clinical efficacy, extending IBD models to clinical scores has been a challenging task. As previously discussed, these scores include highly subjective criteria such as “general wellbeing” and “physician’s rating of disease” proving challenging to “mechanistically link” with disease biology. Even endoscopic scores have been reported to vary between physicians (Kishi et al., 2022). Furthermore, there is limited availability of individual-level patient biopsy data that contain both gut-level biomarkers and clinical score measurements.

To overcome these difficulties, a recent study (Venkatapurapu et al., 2022) has employed a hybrid mechanistic-statistical platform that simultaneously simulates Crohn’s disease progression and incorporates a simple decision tree-based classifier to generate a prediction for SES-CD scores using features such as lesions, biomarkers, and the duration of CD. This platform takes in the patient information, disease profile and treatment history as the input for the responder classifier and generates a long-term response which is then used to identify a matching virtual patient from a virtual population library generated using mechanistic disease modeling. The authors successfully demonstrated that this hybrid method can generate a time series of SES-CD scores along with biomarkers, gut-level cytokines, and immune cell population changes.

This approach of combining a classification algorithm with a mechanistic model is a promising way to extend the prediction because disease activity indices are typically interpreted in categories such as no activity/remission, mild, moderate, and severe. Furthermore, the machine learning component enables learning from literature-reported correlations between biomarkers and clinical scores while the mechanistic model requires a specific data format to be used for calibration. For instance, transcription factors and cytokine levels from isolated tissue biopsies of active IBD patients and healthy controls can be used to correlate with disease severity as potential predictors to be trained and cross-validated in the machine learning model (Li et al., 2017; Olsen et al., 2011).

Additionally, the most notable benefit of the combined approach is that it can be applied to already published models that have been

TABLE 1 Gut biomarkers reported to correlate with clinical scores of IBD.

Correlations	Th1	Th2	Th17	Treg	NK/ NKT	Mac (M1/ M2)	DC	Neutrophil	References
UC endoscopic/ clinical activity			IL-17, IL-21, IL-22, TGFβ	% CD25 ^{hi} FOXP3+ (T cells), TGFβ	IL-17, IL-21, IL-22, TGFβ	IL-8, TGFβ		IL-8	Holmén et al. (2006), Iboshi et al. (2017), Jiang et al. (2014), Zahn et al. (2009)
UCDAI/Mayo	TNFα, IFNγ	% CD4+IL13+FOXP3+ (Th2/Treg)	IL-6	%CD4+IL13+FOXP3+ (Th2/Treg)	IFNγ	TNFα, IL-6, IL-23	IL-6, IL-23, IFNγ	IL-6, IL-23	Allegretti et al. (2023), Li et al. (2017), Olsen et al. (2007), 2011
CDAI	TNFα		IL-6, IL- 17, TGFβ	TGFβ	IL-17, TGFβ	TNFα, IL-6, TGFβ, IL-8	IL-6	IL-6, IL-8	Olsen et al. (2007), 2011; Stallmach et al. (2004)

validated. For instance, Rogers et al. recently published a comprehensive QSP IBD model that can simulate both CD and UC conditions (Rogers et al., 2021a; Rogers et al., 2021b). The model includes important disease biology in the gut and blood involving T cells, macrophages, dendritic cells, and neutrophils along with cytokine release. The model also generates clinical biomarker simulations such as CRP and FCP which are often reported with clinical scores. The authors followed up with additional virtual population analyses by simulating the treatment effects of four different therapeutic targets in CD. They also explored a dual therapy option, anti-TNFα, and anti-IL-12p40, as a potential treatment option, and the simulation predicted that a combination strategy would improve the response compared to mono-therapies (Rogers et al., 2021b). The model code is publicly available to be downloaded in the supporting information section of the Rogers et al. publication.

Given the demonstrated capabilities of the Rogers et al. model, we linked it to a machine learning algorithm to predict clinical scores for both subtypes of IBD and performed a proof-of-principle virtual population analysis.

Clinical score predictions using a combination of a large-scale QSP model and statistical learning approach and example applications

For implementing a clinical score prediction model, we selected the Mayo score (with MES, which is a part of Mayo) for UC and CDAI for CD as output, since they are the most popular metrics in clinical trials. The model was designed to predict the “range of the scores,” defined by clinicians to categorize the severity of the disease. To ensure interpretability, we employed multinomial logistic regression (MLR) in MathWorks (2024), a simple machine-learning classification algorithm, to generate predictions of categorical ranges of clinical scores. Feature selection for the MLR model was guided by the literature evidence between gut biomarkers and the relevant clinical scores (Table 1).

To generate training data, we developed an algorithm that assigns appropriate clinical scores to the published virtual population of the Rogers et al. model based on (1) literature data

of IBD score distribution (Kawashima et al., 2016; Nakamura et al., 2018; Shinzaki et al., 2021) in relation to either FCP or CRP; and (2) simulated levels of tissue biomarkers that had both strong correlations with IBD scores (Holmén et al., 2006) (Table 1) and known mechanistic links with IBD pathology in the literature (Langer et al., 2019). This algorithm computes an overall inflammatory score based on the QSP model-generated steady-state values of relevant tissue biomarkers and serum CRP and FCP levels to match published distributions comparing IBD clinical scores with serum CRP or FCP of individual patients (Kawashima et al., 2016; Nakamura et al., 2018; Shinzaki et al., 2021). Once inflammatory scores were assigned to UC and CD virtual populations, they were used to train the MLR model for the relevant clinical score. An overview of this computational pipeline is described in Figure 1.

The performances of trained models were evaluated using Receiver Operating Characteristic (ROC) curves and sensitivity/specificity metrics (Figure 1; Supplementary Tables S2–S4). ROC curves indicate that the trained models are operating at a consistent level across all classes with moderate to strong predictability specific to the class. Additionally, sensitivity and specificity measures (Supplementary Tables S2–S4) for each MLR model further highlight the strength of the model performance with the average sensitivity/specificity of the MES model at 0.67/0.84, Mayo model at 0.73/0.88, and CDAI model at 0.80/0.89.

Next, exploratory analyses were performed using anti-TNFα (adalimumab), anti-IL-23 (mirikizumab), and the combination. These targets were selected based on the recent study by Feagan et al. (2023) where simultaneous inhibition of TNFα and IL-23 led to significant improvement of therapeutic response in UC patients. The clinical trial data for adalimumab (ECCO, 2024; Puri et al., 2017; Shinzaki et al., 2021) and mirikizumab (Sandborn et al., 2020; Sands et al., 2022) were used to calibrate the model further. Once the clinical score prediction algorithm was built and calibrated, additional analyses were performed to demonstrate whether this hybrid approach can be used to explore (1) mechanistic differences between clinical responders versus non-responders of dual targeting in the UC virtual population; and (2) whether this combination could also be beneficial for treating CD.

For the first part of the exploratory analysis, we generated a UC virtual population (Figure 2A) that matched the patient data

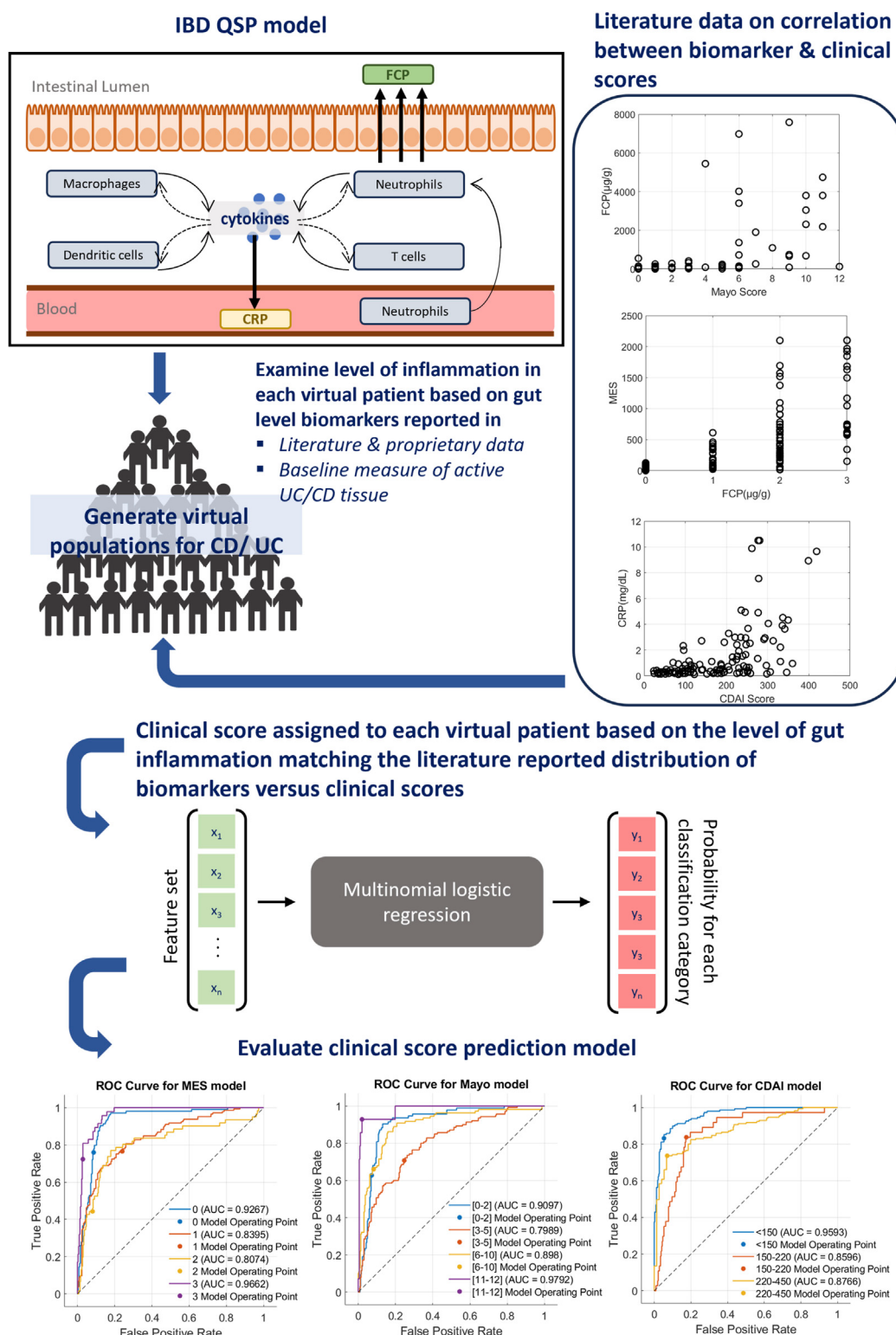


FIGURE 1

Overview of IBD score prediction platform. A clinical prediction algorithm has been integrated with a published mechanistic model of IBD that can simulate both CD and UC. For the interpretability of features, a simple classification algorithm, multinomial logistic regression (MLR), has been selected to build clinical prediction extension. First, the level of inflammation in each virtual patient was computed based on the simulation of relevant tissue biomarkers and serum CRP or FCP. Next, a relevant clinical score was assigned to each virtual patient using the literature reported relationship of clinical scores and CRP or FCP. The actual clinical score distribution data utilized for this process are shown on the right. Using this population, the MLR model was trained to generate clinical score predictions and the performance was evaluated using metrics such as the ROC curve (bottom), sensitivity and specificity measures (Supplementary Tables S2–S4).

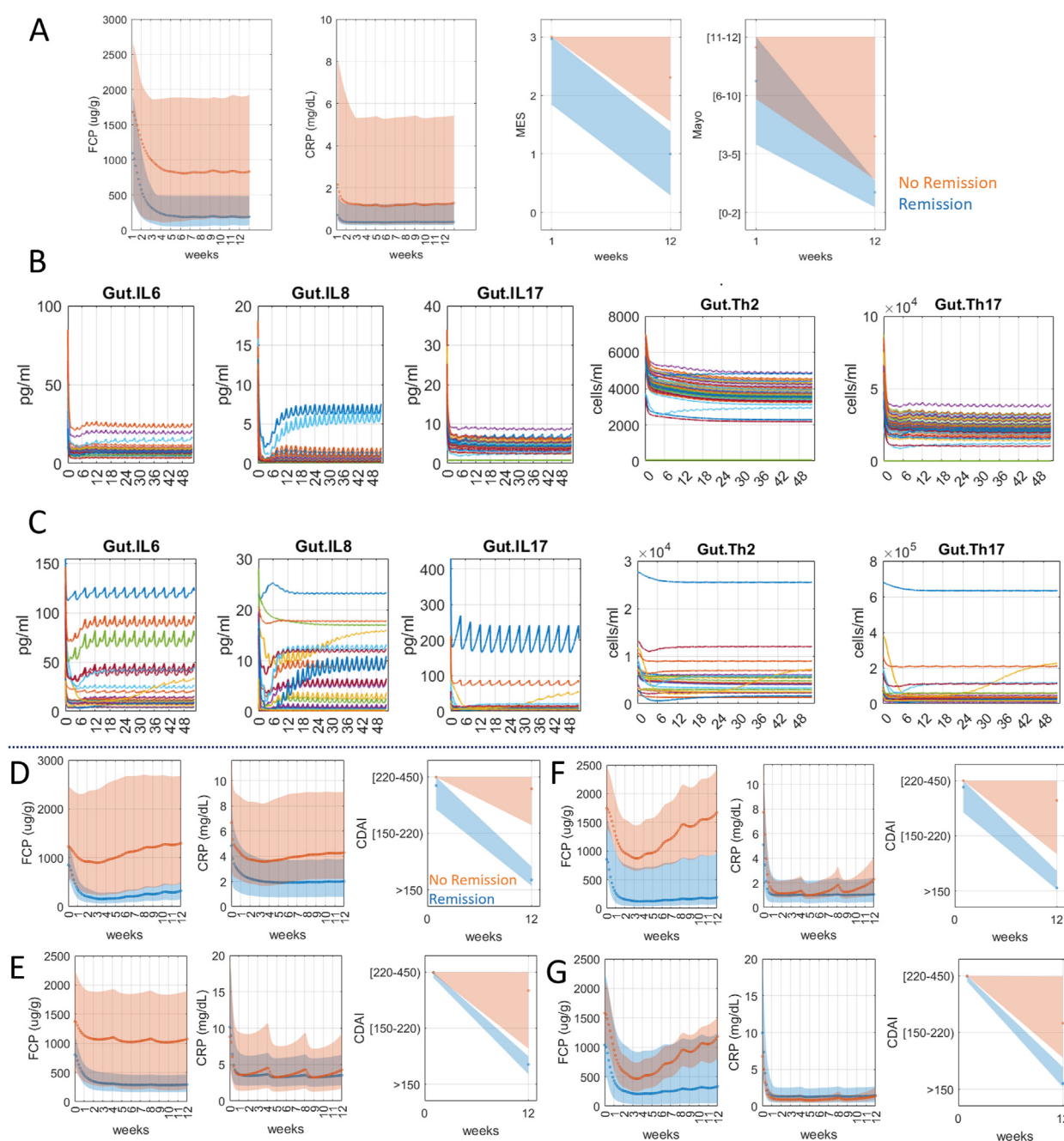


FIGURE 2

Anti-TNF α and anti-IL23 targeting simulations for UC (A–C) and CD (D–G) virtual population. (A) UC virtual population matching the response data of dual targeting of anti-TNF α and anti-IL23 in the VEGA trial (49.3% endoscopic healing and 36.6% clinical remission, MES ≤ 1 and Mayo ≤ 2 , at week 12). Orange shaded area represents the population that did not reach remission in 12 weeks (dark orange dots in CRP/FCP plots represent mean values). The blue shaded area represents the population that did reach remission in 12 weeks (dark blue dots in CRP/FCP plots represent mean values). (B, C) Mechanistic plots of UC virtual population treated with combination therapy. Each line represents an individual virtual patient. The x-axis is on the scale of weeks. Simulation shows dual therapy was not as effective in bringing down IL6, IL8, IL17, Th2 and Th17 in non-responders (C) as in responders (B). (D) CD virtual population matching adalimumab clinical trial data on baseline CDAI (271 ± 56) and treatment response (remission 67% at 12 weeks) based on Shinzaki et al. (2021). (E) CD virtual population matching mirikizumab clinical trial data on baseline CDAI (298 ± 103.7) and treatment response (remission 40.6% at 12 weeks) based on Sands et al. (2022). Both CD virtual populations in (D) and (E) were selected using reported CRP, FCP levels in the trial as well as the reported CDAI distribution at week 12. (F) Virtual CD population matching adalimumab trial data were subjected to hypothetical combination therapy (adalimumab and mirikizumab). The model simulation showed 86% remission at week 12, a significant improvement from the adalimumab trial population (66% remission) (G) Virtual CD population matching mirikizumab trial data were simulated for combination therapy and the result showed 81% remission at week 12, a notable jump from 41%.

published in the VEGA trial (Feagan et al., 2023). This trial showed, that in 12 weeks, 37% of patients treated with combination reached remission while single targeting only reached 21% (anti-IL23, Guselkumab) and 22% (anti-TNF α , Golimumab), respectively. Elaborating on this effort, using mechanistic modeling, we sought to examine the mechanistic differences between the responder population who reached remission in 12 weeks versus the non-responder population (Figures 2B,C). Simulation shows dual targeting was able to significantly decrease IL6, IL8, and IL17 and reduce cell populations of Th2 and Th17 better in responders (Figure 2B) than non-responders (Figure 2C). Additional responder versus non-responder analysis can be performed to identify potential biomarkers of response before treatment.

Next, we evaluated whether anti-TNF α and anti-IL-23 combination could also benefit CD, given the overlapping biology with UC. Because this combination has not been evaluated for CD, we generated two separate CD virtual populations that match the published clinical data for each therapy, adalimumab (Shinzaki et al., 2021) (Figure 2D) and mirikizumab (Sands et al., 2022) (Figure 2E) and evaluated their combination. The results show that, in both virtual populations, combination therapy is predicted to significantly improve the clinical outcome (Figures 2F,G). Simulation shows, for evaluating 12 weeks remission rates, adalimumab and mirikizumab alone were 66% and 41% respectively whereas combination therapy in the same virtual population for both therapies led to 86% and 81%. This is an intriguing result that encourages further testing in the clinic. Thus, a model that can simulate clinical scores enables the exploration of novel drug combinations while providing a mechanistic explanation.

The simulation results show the potential of combining machine learning with mechanistic modeling to predict clinical scores and disease activity indices using publicly available models and data. Furthermore, matching the clinical response rates in UC simulations verifies the clinical score calculation and emphasizes its more general and robust use.

Future directions

A key challenge in validating a predictor of an IBD clinical score is the lack of publicly available individual patient gut biopsy measurements of key cytokines and immune cell activity paired with the actual clinical score. To accommodate for this limitation, we first simulated a validated virtual population from a published model, then selected the virtual patients that matched the FCP or CRP levels of real individual patient data and assigned the associated clinical scores. While the purpose of this manuscript is to showcase a proof-of-principle analysis that would enable extending QSP model prediction to a clinical disease score, an in-depth validation using patient-level biopsy data in the future would strengthen this approach. Another validation step can be to predict various clinical studies with different treatments to validate the MES response predicted by the model. If the prediction achieves enough confidence an analysis of patient endotypes may give important insights to choose the right drug for the right patient.

Another ongoing effort is the CODEX database (Certara, 2024) which is a collection of biomarkers and clinical scores from published

literature sources formatted for validating mechanistic models. Future studies will only enrich the database that can be utilized to bolster the model prediction. On top of these challenges, most clinical scores have empirical and subjective elements that require regression to relate to a mechanistic model. Our QSP model-based approach can leverage population-level distributions of key biomarkers and clinical scores at baseline, induction, and maintenance periods to enrich the UC and CD virtual population improving the accuracy and precision of the QSP model outputs and clinical scores.

Although IBD clinical scores hold the most practical value, significant efforts distinguishing IBD subtypes based on gene expression profiles have been measured at disease onset or the initiation of therapy (Barber et al., 2016). Indeed, integrating omics data sources to inform mechanistic models have been demonstrated in other fields such as oncology and cardiotoxicity (Lazarou et al., 2020; Shim et al., 2023).

Differential gene expression data could serve as inputs to a clinical score predictor; however, such attempts are challenging to validate. Additionally, genetic markers are not regularly measured and may not be robust predictors of response. In addition, such markers are hard to model mechanistically due to the lack of understanding how they link to cell or tissue-level pathophysiology. There has been limited success in other indications such as rheumatoid arthritis (Dennis et al., 2014; Guan et al., 2019; Lewis et al., 2019). In future studies, gene biomarkers may be incorporated into QSP models, alongside cytokine data, to enhance diagnostic or prognostic capabilities.

Conclusion

We presented the contemporary challenges of adding a clinical score predictor to a mechanistic model of IBD illustrating how a combined QSP/machine learning approach could provide a way to overcome those challenges. A recent hybrid approach combined a simple decision tree algorithm with a QSP model to generate SES-CD scores, an endoscopic score of CD, demonstrating the potential of combining machine learning with mechanistic models (Venkatapurapu et al., 2022). Our work expands the clinical prediction for both UC and CD combining a QSP model of IBD with a MLR algorithm to generate relevant clinical scores with interpretable and literature-supported features. Efforts are underway to leverage additional clinical trial data from our CODEX (Certara, 2024) clinical outcomes database to enrich training and testing data sets and improve the accuracy of this approach.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional

requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

Author contributions

JS: Data curation, Formal Analysis, Investigation, Methodology, Software, Writing - original draft. MR: Data curation, Supervision, Formal analysis, Validation, Investigation, Resources, Writing-review and editing. BW: Data curation, Supervision, Formal analysis, Validation, Investigation, Resources, Writing-review and editing. PG: Writing - review and editing. DC: Conceptualization, Data curation, Investigation, Methodology, Software, Writing - review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

We would like to acknowledge Jaap Mandema (Certara) for their expert input and support.

References

- Abraham, C., and Cho, J. H. (2009). Inflammatory bowel disease. *N. Engl. J. Med.* 361 (21), 2066–2078. doi:10.1056/NEJMRA0804647
- Ahluwalia, B., Moraes, L., Magnusson, M. K., and Öhman, L. (2018). Immunopathogenesis of inflammatory bowel disease and mechanisms of biological therapies. *Scand. J. Gastroenterology* 53 (4), 379–389. doi:10.1080/00365521.2018.1447597
- Allegretti, J. R., Mitsialis, V., Canavan, J. B., Hamilton, M., Barends, J., Carrellas, M., et al. (2023). Low-dose interleukin 2 for the treatment of moderate to severe ulcerative colitis. *Gastroenterology* 165 (2), 492–495.e2. doi:10.1053/j.GASTRO.2023.03.230
- Balbas-Martinez, V., Ruiz-Cerdá, L., Irurzun-Arana, I., González-García, I., Vermeulen, A., Gómez-Mantilla, J. D., et al. (2018). A systems pharmacology model for inflammatory bowel disease. *PLOS ONE* 13 (3), e0192949. doi:10.1371/JOURNAL.PONE.0192949
- Barber, G. E., Yajnik, V., Khalili, H., Giallourakis, C., Garber, J., Xavier, R., et al. (2016). Genetic markers predict primary non-response and durable response to anti-TNF biologic therapies in Crohn's disease. *Am. J. Gastroenterology* 111 (12), 1816–1822. doi:10.1038/AJG.2016.408
- CADTH (2016). Clinical review report: adalimumab (humira). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK539025/>.
- Cai, Z., Wang, S., and Li, J. (2021). Treatment of inflammatory bowel disease: a comprehensive review. *Front. Med.* 8, 765474. doi:10.3389/FMED.2021.765474
- Certara (2024). CODEx and clinical outcomes databases | quantifying data. Available at: <https://www.certara.com/data-and-informatics/codex-clinical-trial-outcomes-databases/> (Accessed June 22, 2024).
- Dai, C., Huang, Y. H., and Jiang, M. (2023). Combination therapy in inflammatory bowel disease: current evidence and perspectives. *Int. Immunopharmacol.* 114, 109545. doi:10.1016/J.INTIMP.2022.109545
- Dawoud, N., Roy, L., Elmoursi, A., Elkammar, M., Eltaher, M., Barrett, T., et al. (2022). Combination ustekinumab and vedolizumab therapy for refractory inflammatory bowel disease: a single institution case series. *Gastroenterology* 162 (3), S103–S104. doi:10.1053/j.gastro.2021.12.215
- Dennis, G., Holweg, C. T. J., Kummerfeld, S. K., Choy, D. F., Setiadi, A. F., Hackney, J. A., et al. (2014). Synovial phenotypes in rheumatoid arthritis correlate with response to biologic therapeutics. *Arthritis Res. Ther.* 16 (2), R90–R18. doi:10.1186/AR4555
- ECCO (2024). DOP24 Patient-reported health-related quality-of-life outcomes with vedolizumab vs. adalimumab treatment of ulcerative colitis: results of the VARSITY trial. Available at: <https://www.ecco-ibd.eu/publications/congress-abstracts/item/dop24-patient-reported-health-related-quality-of-life-outcomes-with-vedolizumab-vs-adalimumab-treatment-of-ulcerative-colitis-results-of-the-verity-trial.html> (Accessed April 20, 2024).
- Falvey, J. D., Hoskin, T., Meijer, B., Ashcroft, A., Walmsley, R., Day, A. S., et al. (2015). Disease activity assessment in IBD: clinical indices and biomarkers fail to predict endoscopic remission. *Inflamm. Bowel Dis.* 21 (4), 824–831. doi:10.1097/MIB.0000000000000341
- Feagan, B. G., Sands, B. E., Sandborn, W. J., Germinaro, M., Vetter, M., Shao, J., et al. (2023). Guselkumab plus golimumab combination therapy versus guselkumab or golimumab monotherapy in patients with ulcerative colitis (VEGA): a randomised, double-blind, controlled, phase 2, proof-of-concept trial. *Lancet Gastroenterology Hepatology* 8 (4), 307–320. doi:10.1016/S2468-1253(22)00427-7
- Fendt, R., Karamitsou, V., Rehberg, M., Flechsenhar, K., and Wagenhuber, B. (2024). A virtual ulcerative colitis patient population to explore combination treatments in silico clinical trials. PAGE2024: drug/Disease Modelling Abstract. Available at: <https://www.page-meeting.org/default.asp?abstract=10905>.
- Freeman, H. J. (2008). Use of the Crohn's disease activity index in clinical trials of biological agents. *World J. Gastroenterology WJG* 14 (26), 4127–4130. doi:10.3748/WJG.14.4127
- Gajendran, M., Loganathan, P., Catinella, A. P., and Hashash, J. G. (2018). A comprehensive review and update on Crohn's disease. *Disease-a-Month* 64 (2), 20–57. doi:10.1016/J.DISAMONTH.2017.07.001
- Gajendran, M., Loganathan, P., Jimenez, G., Catinella, A. P., Ng, N., Umapathy, C., et al. (2019). A comprehensive review and update on ulcerative colitis. *Disease-a-Month* 65 (12), 100851. doi:10.1016/J.DISAMONTH.2019.02.004
- Goessens, L., Colomel, J. F., Outtier, A., Ferrante, M., Sabino, J., Judge, C., et al. (2021). Safety and efficacy of combining biologics or small molecules for inflammatory bowel disease or immune-mediated inflammatory diseases: a European retrospective observational study. *UEG J.* 9 (10), 1136–1147. doi:10.1002/UEG2.12170
- Guan, Y., Zhang, H., Quang, D., Wang, Z., Parker, S. C. J., Pappas, D. A., et al. (2019). Machine learning to predict anti-tumor necrosis factor drug responses of rheumatoid arthritis patients by integrating clinical and genetic markers. *Arthritis Rheumatology* 71 (12), 1987–1996. doi:10.1002/ART.41056

Conflict of interest

MR and BW are employees of Sanofi and may hold stock and/or stock options in the company. PG, DC, and JS are employees of Certara and may hold stock in the company. PG, CD, and JS have served as consultants for Sanofi for the development of the model and algorithm that were utilized in this study. Certara and Sanofi hold no stake in the molecule adalimumab, clinically approved for the treatment of Ulcerative Colitis and Crohn's Disease, nor mirikizumab, clinically approved for the treatment of Ulcerative Colitis.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphar.2025.1479666/full#supplementary-material>

- Holmén, N., Lundgren, A., Lundin, S., Bergin, A. M., Rudin, A., Sjövall, H., et al. (2006). Functional CD4+CD25high regulatory T cells are enriched in the colonic mucosa of patients with active ulcerative colitis and increase with disease activity. *Inflamm. Bowel Dis.* 12 (6), 447–456. doi:10.1097/00054725-200606000-00003
- Iboshi, Y., Nakamura, K., Fukaura, K., Iwasa, T., Ogino, H., Sumida, Y., et al. (2017). Increased IL-17A/IL-17F expression ratio represents the key mucosal T helper/regulatory cell-related gene signature paralleling disease activity in ulcerative colitis. *J. Gastroenterology* 52 (3), 315–326. doi:10.1007/S00535-016-1221-1
- Jiang, W., Su, J., Zhang, X., Cheng, X., Zhou, J., Shi, R., et al. (2014). Elevated levels of Th17 cells and Th17-related cytokines are associated with disease activity in patients with inflammatory bowel disease. *Inflamm. Res.* 63 (11), 943–950. doi:10.1007/S00011-014-0768-7
- Kawashima, K., Ishihara, S., Yuki, T., Fukuba, N., Oshima, N., Kazumori, H., et al. (2016). Fecal calprotectin level correlated with both endoscopic severity and disease extent in ulcerative colitis. *BMC Gastroenterol.* 16 (1), 47. doi:10.1186/S12876-016-0462-Z
- Kilian, C., Ulrich, H., Zouboulis, V. A., Sprezyna, P., Schreiber, J., Landsberger, T., et al. (2024). Longitudinal single-cell data informs deterministic modelling of inflammatory bowel disease. *Npj Syst. Biol. Appl.* 10 (1), 69–9. doi:10.1038/s41540-024-00395-9
- Kim, E., Kang, Y., Lee, M. J., Park, Y. N., and Koh, H. (2013). Life-threatening lower gastrointestinal hemorrhage in pediatric Crohn's disease. *Pediatr. Gastroenterology, Hepatology and Nutr.* 16 (1), 53–60. doi:10.5223/PGHN.2013.16.1.53
- Kim, K. O. (2022). Endoscopic activity in inflammatory bowel disease: clinical significance and application in practice. *Clin. Endosc.* 55 (4), 480–488. doi:10.5946/CE.2022.108
- Kishi, M., Hirai, F., Takatsu, N., Hisabe, T., Takada, Y., Beppu, T., et al. (2022). A review on the current status and definitions of activity indices in inflammatory bowel disease: how to use indices for precise evaluation. *J. Gastroenterology* 57 (4), 246–266. doi:10.1007/S00535-022-01862-Y
- Kobayashi, T., and Hibi, T. (2023). Improving IBD outcomes in the era of many treatment options. *Nat. Rev. Gastroenterology and Hepatology* 20 (2), 79–80. doi:10.1038/s41575-022-00738-z
- Langer, V., Vivi, E., Regensburger, D., Winkler, T. H., Waldner, M. J., Rath, T., et al. (2019). IFN- γ drives inflammatory bowel disease pathogenesis through VE-cadherin-directed vascular barrier disruption. *J. Clin. Investigation* 129 (11), 4691–4707. doi:10.1172/JCI124884
- Lazarou, G., Chelliah, V., Small, B. G., Walker, M., van der Graaf, P. H., and Kierzek, A. M. (2020). Integration of omics data sources to inform mechanistic modeling of immune-oncology therapies: a tutorial for clinical pharmacologists. *Clin. Pharmacol. and Ther.* 107 (4), 858–870. doi:10.1002/CPT.1786
- Lewis, J. D., Chuai, S., Nessel, L., Lichtenstein, G. R., Aberra, F. N., and Ellenberg, J. H. (2008). Use of the non-invasive components of the mayo score to assess clinical response in ulcerative colitis. *Inflamm. Bowel Dis.* 14 (12), 1660–1666. doi:10.1002/IBD.20520
- Lewis, M. J., Barnes, M. R., Blighe, K., Goldmann, K., Rana, S., Hackney, J. A., et al. (2019). Molecular portraits of early rheumatoid arthritis identify clinical and treatment response phenotypes. *Cell Rep.* 28 (9), 2455–2470. doi:10.1016/j.celrep.2019.07.091
- Li, J., Ueno, A., Iacucci, M., Fort Gasia, M., Jijon, H. B., Panaccione, R., et al. (2017). Crossover subsets of CD4+ T lymphocytes in the intestinal lamina propria of patients with Crohn's disease and ulcerative colitis. *Dig. Dis. Sci.* 62 (9), 2357–2368. doi:10.1007/S10620-017-4596-9
- Liu, J. K., and Lichtenstein, G. R. (2012). "Inflammatory bowel disease," in *Clinical gastrointestinal endoscopy*. Second Edition, 243–264. doi:10.1016/B978-1-4377-1529-3.00021-X
- Lo, W. C., Arsenescu, R. I., and Friedman, A. (2013). Mathematical model of the roles of T cells in inflammatory bowel disease. *Bull. Math. Biol.* 75 (9), 1417–1433. doi:10.1007/s11538-013-9853-2
- MathWorks (2024). Mnrfit - multinomial logistic regression - MATLAB. Available at: <https://www.mathworks.com/help/stats/mnrfit.html> (Accessed November 3, 2024).
- Nakamura, S., Imaeda, H., Nishikawa, H., Imuro, M., Matsuura, M., Oka, H., et al. (2018). Usefulness of fecal calprotectin by monoclonal antibody testing in adult Japanese with inflammatory bowel diseases: a prospective multicenter study. *Intestinal Res.* 16 (4), 554–562. doi:10.5217/IR.2018.00027
- Olsen, T., Goll, R., Cui, G., Husebekk, A., Vonen, B., Birketvedt, G. S., et al. (2007). Tissue levels of tumor necrosis factor- α correlates with grade of inflammation in untreated ulcerative colitis. *Scand. J. Gastroenterology* 42 (11), 1312–1320. doi:10.1080/00365520701409035
- Olsen, T., Rismo, R., Cui, G., Goll, R., Christiansen, I., and Florholmen, J. (2011). TH1 and TH17 interactions in untreated inflamed mucosa of inflammatory bowel disease, and their potential to mediate the inflammation. *Cytokine* 56 (3), 633–640. doi:10.1016/J.CYTO.2011.08.036
- Pinton, P. (2022). Computational models in inflammatory bowel disease. *Clin. Transl. Sci.* 15 (4), 824–830. doi:10.1111/CTS.13228
- Pinton, P. (2023). Impact of artificial intelligence on prognosis, shared decision-making, and precision medicine for patients with inflammatory bowel disease: a perspective and expert opinion. *Ann. Med.* 55 (2), 2300670. doi:10.1080/07853890.2023.2300670
- Puri, A., Niewiarowski, A., Arai, Y., Nomura, H., Baird, M., Dalrymple, I., et al. (2017). Pharmacokinetics, safety, tolerability and immunogenicity of FKB327, a new biosimilar medicine of adalimumab/Humira, in healthy subjects. *Br. J. Clin. Pharmacol.* 83 (7), 1405–1415. doi:10.1111/BCP.13245
- Ramos, G. P., and Papadakis, K. A. (2019). Mechanisms of disease: inflammatory bowel diseases. *Mayo Clin. Proc.* 94, 155–165. doi:10.1016/j.mayocp.2018.09.013
- Rogers, K. V., Martin, S. W., Bhattacharya, I., Singh, R. S. P., and Nayak, S. (2021a). A dynamic quantitative systems pharmacology model of inflammatory bowel disease: Part 1 – model framework. *Clin. Transl. Sci.* 14 (1), 239–248. doi:10.1111/CTS.12849
- Rogers, K. V., Martin, S. W., Bhattacharya, I., Singh, R. S. P., and Nayak, S. (2021b). A dynamic quantitative systems pharmacology model of inflammatory bowel disease: Part 2 – application to current therapies in Crohn's disease. *Clin. Transl. Sci.* 14 (1), 249–259. doi:10.1111/CTS.12850
- Saez, A., Herrero-Fernandez, B., Gomez-Bris, R., Sánchez-Martínez, H., and Gonzalez-Granado, J. M. (2023). Pathophysiology of inflammatory bowel disease: innate immune system. *Int. J. Mol. Sci.* 24 (2), 1526. doi:10.3390/IJMS24021526
- Sandborn, W. J., Ferrante, M., Bhandari, B. R., Berliba, E., Feagan, B. G., Hibi, T., et al. (2020). Efficacy and safety of mirikizumab in a randomized phase 2 study of patients with ulcerative colitis. *Gastroenterology* 158 (3), 537–549. doi:10.1053/j.gastro.2019.08.043
- Sands, B. E., Peyrin-Biroulet, L., Kierkus, J., Higgins, P. D. R., Fischer, M., Jairath, V., et al. (2022). Efficacy and safety of mirikizumab in a randomized phase 2 study of patients with Crohn's disease. *Gastroenterology* 162 (2), 495–508. doi:10.1053/J.GASTRO.2021.10.050
- Sartor, R. B. (2006). Mechanisms of Disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nat. Clin. Pract. Gastroenterology and Hepatology* 3 (7), 390–407. doi:10.1038/ncpgasthep0528
- Schroeder, K. W., Tremaine, W. J., and Ilstrup, D. M. (1987). Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N. Engl. J. Med.* 317 (26), 1625–1629. doi:10.1056/NEJM198712243172603
- Shim, J. V., Xiong, Y., Dhanan, P., Dariolli, R., Azeloglu, E. U., Hu, B., et al. (2023). Predicting individual-specific cardiotoxicity responses induced by tyrosine kinase inhibitors. *Front. Pharmacol.* 14, 1158222. doi:10.3389/fphar.2023.1158222
- Shinzaki, S., Matsuoka, K., Tanaka, H., Takeshima, F., Kato, S., Torisu, T., et al. (2021). Leucine-rich alpha-2 glycoprotein is a potential biomarker to monitor disease activity in inflammatory bowel disease receiving adalimumab: PLANET study. *J. Gastroenterology* 56 (6), 560–569. doi:10.1007/S00535-021-01793-0
- Stallmach, A., Giese, T., Schmidt, C., Ludwig, B., Mueller-Molaiian, I., and Meuer, S. C. (2004). Cytokine/chemokine transcript profiles reflect mucosal inflammation in Crohn's disease. *Int. J. Colorectal Dis.* 19 (4), 308–315. doi:10.1007/S00384-003-0554-4
- Stübler, S., Kloft, C., and Huisinga, W. (2023). Cell-level systems biology model to study inflammatory bowel diseases and their treatment options. *CPT Pharmacometrics and Syst. Pharmacol.* 12 (5), 690–705. doi:10.1002/PSP.4.12932
- Sultan, K. S., Berkowitz, J. C., and Khan, S. (2017). Combination therapy for inflammatory bowel disease. *World J. Gastrointest. Pharmacol. Ther.* 8 (2), 103–113. doi:10.4292/WJGPT.V8.I2.103
- Venkatapurapu, S. P., Iwakiri, R., Udagawa, E., Patidar, N., Qi, Z., Takayama, R., et al. (2022). A computational platform integrating a mechanistic model of Crohn's disease for predicting temporal progression of mucosal damage and healing. *Adv. Ther.* 39 (7), 3225–3247. doi:10.1007/S12325-022-02144-Y
- Wagatsuma, K., Yokoyama, Y., and Nakase, H. (2021). Role of biomarkers in the diagnosis and treatment of inflammatory bowel disease. *Life Basel, Switz.* 11 (12), 1375. doi:10.3390/LIFE11121375
- Wendelsdorf, K., Bassaganya-Riera, J., Hontecillas, R., and Eubank, S. (2010). Model of colonic inflammation: immune modulatory mechanisms in inflammatory bowel disease. *J. Theor. Biol.* 264 (4), 1225–1239. doi:10.1016/J.JTBI.2010.03.027
- Wetwittayakhang, P., and Lakatos, P. L. (2024). Current evidence for combined targeted therapy for the treatment of inflammatory bowel disease. *J. Can. Assoc. Gastroenterology* 7 (1), 22–29. doi:10.1093/JCAG/GWAD032
- Whittaker, D. G., Shaliban, A., Roy, M., Packirisamy, P., Rengaswamy, M., Damian, V., et al. (2024). Leveraging *in vitro* data from novel drug candidates to prioritize antibody combinations in autoimmune disease using a QSP model of IBD. *PAGE2024 Drug/Disease Model. Abstr.* Available at: <https://www.page-meeting.org/default.asp?abstract=10876>.
- Yu, B., Jin, S., Han, J., Xu, J., Zhang, S., Li, Y., et al. (2024). Quantitative evaluation of the time-course and efficacy of targeted agents for ulcerative colitis. *Front. Pharmacol.* 15, 1399963. doi:10.3389/fphar.2024.1399963
- Zahn, A., Giese, T., Karner, M., Braun, A., Hinz, U., Stremmel, W., et al. (2009). Transcript levels of different cytokines and chemokines correlate with clinical and endoscopic activity in ulcerative colitis. *BMC Gastroenterol.* 9, 13. doi:10.1186/1471-230X-9-13