

Utility of Machine Learning to Characterize Gut Microbiota Dysbiosis and Its Clinical Implications in Inflammatory Bowel Disease

June-Young Lee^{1,2,*}, Dong Hyun Kim^{3,*}, Jee-Won Choi⁴, Minho Shong⁵, Chang Kyun Lee⁶

¹InnoCORE AI-CRED Institute, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea; ²Thor Therapeutics Inc., Daejeon, South Korea; ³Division of Gastroenterology, Department of Internal Medicine, Chonnam National University Hospital and Medical School, Gwangju, South Korea; ⁴Department of Biology, Kyung Hee University, Seoul, South Korea; ⁵Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology, Daejeon, South Korea; ⁶Department of Gastroenterology, Center for Crohn's and Colitis, Kyung Hee University Hospital, Kyung Hee University College of Medicine, Seoul, South Korea

*These authors contributed equally to this work

Correspondence: Chang Kyun Lee, Department of Gastroenterology, Center for Crohn's and Colitis, Kyung Hee University Hospital, Kyung Hee University College of Medicine, Seoul, South Korea, Tel +82-2-958-8238, Email changkyun.lee@khu.ac.kr

Abstract: Inflammatory bowel disease (IBD) arises from complex interactions among host genetics, immune dysregulation, environmental factors, and the gut microbiome. Numerous studies have demonstrated alterations in microbial composition and function, including reduced diversity and changes in metabolic pathways. Traditional biostatistical approaches, such as differential abundance analysis, have advanced our understanding but are remain limited in handling nonlinear and high-dimensional data. Machine learning (ML) complements these methods by integrating heterogeneous datasets and uncovering hidden patterns that improve classification and predictive accuracy. In IBD, delayed diagnosis and the lack of reliable biomarkers highlight the need for computational tools that can translate complex microbiome data into clinically actionable insights. ML and deep learning (DL) have expanded analytical capabilities, enabling disease classification, subtype differentiation, and prediction of therapeutic responses. This review provides an integrative perspective on how ML and DL are reshaping microbiome-based IBD research, summarizing their strengths, limitations, and essential considerations for clinical translation. Future progress will depend on standardized microbiome assays, rigorous benchmarking, and the integration of multi-omics data to elucidate host-microbe interactions. With these advancements, ML- and DL-based approaches may offer precise diagnostics and personalized treatment strategies, transforming microbiome research into practical tools for IBD care.

Keywords: microbiome, machine-learning, inflammatory bowel disease, dysbiosis, artificial intelligence

Introduction

Inflammatory bowel diseases (IBD), including Crohn's disease (CD) and ulcerative colitis (UC), result from interactions between the host, environmental factors, and gut microbiome.^{1,2} A common feature of IBD is microbial community alteration, characterized by reduced microbial diversity, depletion of short-chain fatty acid-producing bacteria, enrichment of opportunistic taxa, and disrupted metabolic pathways that change with disease activity and therapeutic interventions.³⁻⁵ Recent longitudinal and multi-omics studies have shown that these changes in compositional and functional shifts are closely related to host transcriptomic, proteomic, and metabolomic signatures, underscoring the centrality of host-microbe interactions in disease pathogenesis.⁶⁻⁸

Despite these advances, significant clinical unmet needs remain. The diagnosis of IBD is often delayed because symptoms overlap with those of other gastrointestinal disorders,^{9,10} and current biomarkers such as C-reactive protein and fecal calprotectin lack disease specificity and prognostic utility.¹¹⁻¹³ Reliable microbial or molecular biomarkers capable of predicting treatment response or relapse risk are still lacking, highlighting the need for integrative analytic

frameworks that can bridge the gap between discovery and clinical application.^{14,15} Traditional statistical analyses, including univariate (Linear discriminant analysis Effect Size, LEfSe) or multivariate regression models (Microbiome Multivariable Association with Linear Models, MaAsLin2), have been instrumental in identifying differentially abundant taxa and host factors.^{16,17} However, these approaches are limited in capturing nonlinear interactions and complex dependencies within high-dimensional, multi-omics datasets.^{18,19}

These datasets are intrinsically high-dimensional, sparse, and heterogeneous across cohorts, sampling sites, sequencing platforms, and geographies, with outcomes that can vary considerably depending on the analytical methods applied. Longitudinal and cross-cohort studies have revealed significant inter-individual variability and nonlinear dependencies among microbes, host factors, and metabolites, which makes finding biomarkers more complicated.^{7,20} Additionally, large-scale integration of metagenomic and metabolomic data has shown both the potential to identify disease-related features and the ongoing challenge of aligning results across studies. This emphasizes the need for strong computational methods to handle this complexity.^{6,21}

In this context, machine learning (ML) has emerged as a robust framework for characterizing gut microbial changes in IBD.²² ML approaches are particularly well-suited for (i) extracting predictive patterns from high-dimensional data, (ii) integrating heterogeneous inputs spanning microbial taxa, functional genes, host omics, and clinical metadata, and (iii) delivering measurable classification performance for diagnostic, prognostic, and therapeutic stratification.²³ Applications in IBD cohorts have demonstrated that ML models outperform traditional statistical methods in discriminating cases from controls, differentiating subtypes, and predicting treatment responses.^{4,24} Simultaneously, benchmarking studies and recent state-of-the-art reviews have emphasized critical challenges—including external validation, model calibration, interpretability, and transparent reporting—that must be addressed before ML-driven microbiome biomarkers can be reliably translated into clinical settings.^{18,23,25}

Building on recent reviews that highlighted the importance of methodological robustness, reproducibility, and the need for clinical translation in AI-driven microbiome research,^{26,27} this review extends these perspectives by incorporating frameworks such as time-series and causal inference modeling, multi-omics data integration, and foundation-model approaches to enhance translational relevance. We provide a comprehensive overview of recent applications of ML for characterizing gut microbiome alterations in IBD, emphasizing commonly used algorithms, analytical pipelines, and model evaluation strategies. Finally, we highlight the diagnostic and prognostic potential of ML-based approaches and discuss their strengths, limitations, and key considerations in translating microbiome-derived insights into clinically applicable tools for precision medicine.

Microbiome Dysbiosis in IBD: From Traditional Studies to Multi-Omics Analyses

Early studies on gut microbiome in IBD primarily relied on 16S rRNA gene sequencing, which enabled cost-effective profiling of microbial communities at the genus or family level. These studies provided the first evidence of disease-associated dysbiosis characterized by reduced alpha diversity, enrichment of pathobionts, and depletion of commensals.^{3,5} Initial analytical approaches focused mainly on univariate comparisons of relative abundances, typically using methods such as LEfSe.¹⁷ Ma et al applied LEfSe analysis to compare patients with early-stage CD, advanced CD, and healthy controls, identifying enrichment of *Parabacteroides* and *Lachnospiraceae incertae sedis* in early CD, expansion of *Escherichia-Shigella* and *Proteus* in advanced CD, and preservation of short-chain fatty acid-producing taxa such as *Roseburia* and *Butyrivibrio* in healthy controls.²⁸ These findings demonstrate how LEfSe highlights disease-stage-specific microbial signatures while underscoring its reliance on univariate contrasts.

However, subsequent microbiome analyses have increasingly adopted multivariate analyses that incorporated meta-data-based adjustments and advanced statistical modeling. Tools, such as MaAsLin2¹⁶/MaAsLin3²⁹ and ANCOM-BC³⁰/ANCOM-BC2,³¹ further enhance feature selection by integrating clinical metadata and complex covariates. Chen et al used MaAsLin2 to identify taxa whose relative abundances differed in quiescent CD (CD-R), independent of active inflammation and environmental or genetic confounders.³² They found that *Faecalibacterium*, *Dorea*, and *Fusicatenibacter* were significantly decreased in patients with CD-R compared to healthy first-degree relatives and non-

relative healthy controls. Rosso et al applied ANCOM-BC to fecal bacterial profiles of patients with IBD (both UC and CD) vs non-IBD controls in Buenos Aires.³³ Among the bacterial genera that were differentially abundant in ANCOM-BC, *Bifidobacterium* was enriched in patients with UC compared to non-IBD controls, whereas in patients with CD, *Bifidobacterium*, *Bacteroides*, *Lactobacillus*, and *Faecalibacterium* were identified as differentially abundant taxa. Together, these approaches reduced the analytical bias and enabled a more reproducible investigation of disease-specific taxa, marking a substantial advancement from exploratory compositional analyses to clinically grounded biomarker discovery. In addition, other methods such as ALDEx2³⁴ and LinDA,³⁵ although not discussed in detail in this review, represent parallel efforts to further mitigate the biases inherent to differential abundance analysis.

The advent of high-throughput shotgun metagenomic sequencing has provided a technological leap, offering strain-level resolution and direct access to microbial functional repertoires, including genes and pathways relevant to host metabolism and immunity.³⁶ In a recent multibiome study, Akiyama et al applied shotgun metagenomic sequencing with MaAsLin2 across Japanese and external validation cohorts to characterize microbial signatures in IBD.³⁷ Both UC and CD are associated with the depletion of short-chain fatty acid-producing bacteria and the enrichment of *Enterococcus faecium* and *Bifidobacterium* species. Notably, *Escherichia coli* specifically increased only in patients with CD, highlighting its potential role as a disease-specific microbial marker.

Importantly, gut microbiome dysbiosis in patients with IBD is not limited to bacteria. Studies on the gut virome^{38,39} have reported an increased abundance of bacteriophages (eg, Caudovirales) and decreased abundance of eukaryotic viruses, while mycobiome studies have identified altered fungal-bacterial interactions and increased abundance of *Candida* species.⁴⁰ Beyond the microbial taxa themselves, metabolomic and proteomic analyses revealed disease-associated alterations in bile acids, short-chain fatty acids, and amino acid derivatives, providing functional readouts that link microbial activity to host immune and epithelial responses.^{41,42} Together, these multi-omics approaches provide a more comprehensive understanding of microbial changes and its systemic effects.

However, the integration of multiple data layers has resulted in significant analytical challenges. Microbiome and multi-omics datasets are intrinsically high-dimensional, sparse, and heterogeneous with marked variability across cohorts, sequencing platforms, and sample types.^{18,19,43} Traditional univariate and linear statistical approaches are often insufficient for capturing nonlinear dependencies and cross-modal interactions. Although analytical tools such as LefSe and MaAsLin2 are widely used for differential abundance and association testing, they are inherently sensitive to compositional bias, differences in sequencing depth, and batch effects.^{19,44} LefSe operates on relative abundances and may generate spurious associations when sample library sizes vary, whereas MaAsLin2 assumes linear relationships and is affected by zero inflation and normalization artifacts. These issues have been repeatedly highlighted in recent benchmarking studies, emphasizing the need for compositional data-aware methods and rigorous normalization to ensure cross-cohort reproducibility and biological validity.⁴⁵ Collectively, these challenges have motivated the adoption of ML and related artificial intelligence (AI) approaches, which are particularly well-suited for recognizing complex patterns in high-dimensional data and enabling multimodal integration.^{46,47} Leveraging these advanced computational frameworks, researchers aim to move beyond traditional analyses toward more reproducible, predictive, and clinically interpretable models of gut microbiome dysbiosis in IBD.

Machine Learning Applications in IBD Microbiome Research

As the complexity of microbiome data in IBD has become more apparent, ML approaches have been increasingly adopted to address the limitations of conventional statistical analyses.^{18,48} Unlike traditional differential abundance methods, ML algorithms can handle high-dimensional and heterogeneous data and discover subtle nonlinear patterns that may better capture disease-associated microbial biomarkers.^{49,50} These shifts have established ML as a key tool for developing diagnostic and prognostic models of IBD.^{51,52}

Disease Classification and Subtype Diagnosis

One of the earliest and most widely explored applications of ML in microbiome research was disease diagnosis. Classifier models trained on microbial abundance profiles or functional gene features have been used to discriminate patients with IBD from healthy controls,²⁴ distinguish patients with CD from those with UC,⁵⁰ and predict clinical outcomes such as flare

risk, therapeutic response, or remission status.^{52,53} In many cases, ML classifiers have achieved significant performance as measured via receiver operating characteristic (ROC) curves, the area under the curves (AUCs), sensitivity, and specificity, often surpassing the predictive capacity of conventional biomarkers such as fecal calprotectin.²⁴ For example, Liñares-Blanco et al applied a panel of supervised learning algorithms to fecal microbiome data and developed a robust signature capable of classifying not only IBD versus controls but also differentiating CD from UC.⁵⁰ Their random forest (RF)-based model achieved high AUC values across cross-validation, and feature selection identified a reproducible set of bacterial taxa that contributed most strongly to classification. Similarly, Park et al reported that a fecal microbiome-based ML model accurately discriminated between CD and UC, further confirming that microbial signatures can serve as reliable classifiers of IBD subtypes.⁵¹ Recent large-scale studies have underscored the diagnostic power of microbiome-informed ML. Zheng et al's multi-ethnic cohort analysis demonstrated that an ML classifier trained on microbial taxonomic features achieved AUCs exceeding 0.7–0.9, outperforming conventional biomarkers such as fecal calprotectin in identifying both UC and CD.²⁴ Similarly, Kim et al constructed an ML model using Korean fecal metagenomes that reliably differentiated patients with IBD from healthy patients, emphasizing the generalizability of microbiome signatures across populations.⁵⁴ Together, these studies highlight that microbiome-based ML frameworks may evolve into noninvasive diagnostic tools capable of complementing or reducing the need for invasive endoscopy.

Biomarker Discovery and Algorithmic Approaches

In addition to its classification capabilities, ML plays a key role in the discovery of disease-specific features and biomarkers.^{55,56} By ranking microbial taxa, genes, or metabolites according to their predictive importance, algorithms not only improve model interpretability but also generate hypotheses about how microbes might contribute to diseases. This capability has enabled the development of candidate biomarker panels that integrate microbial and metabolic features with clinical metadata, thereby providing a more comprehensive view of IBD pathophysiology. Among the various algorithms employed, RF⁵⁷ is the most widely used in microbiome studies because of its robustness to noise, sparsity tolerance, and straightforward feature importance readouts.⁵⁸ Knights et al highlighted its stability and performance compared to traditional statistical approaches.⁵⁹ Kim et al further applied RF-based feature selection to identify microbial markers associated with IBD, distinguishing taxa such as *Veillonella* and *Escherichia-Shigella* between patients with UC and those with CD.⁵⁴ Manandhar et al also applied RF classifiers trained on gut microbiome features to diagnose IBD and differentiate CD from UC using the American Gut Project data comprising 729 patients with IBD and 700 healthy patients.⁴⁷ With 50 LEfSe-selected taxa, the RF model achieved a test AUCs of > 0.80 for IBD versus controls; using the top 500 high-variance operational taxonomic units (OTUs) further improved performance to an AUCs > 0.82. For subtype classification (331 CD vs 141 UC), RF models trained on either differential taxa or high-variance OTUs attained test AUCs > 0.90, underscoring their strong discriminatory power and potential clinical utility.

Although RF remains the most commonly used approach, other algorithms, including support vector machines (SVMs) and boosting frameworks (eg, XGBoost and AdaBoost), are also widely applied.^{60,61} In an extensive multiomics study of CD location subtypes, Gonzalez et al systematically evaluated ten different algorithms (including RFs, extra trees, decision trees, SVMs, multilayer perceptron classifiers, voting classifiers, naïve Bayes, k-nearest neighbors, logistic regression, and AdaBoost) for their ability to classify disease subtypes. The authors identified the best-performing models and applied them to downstream analyses, revealing distinct associations according to disease location. Specifically, colonic CD exhibited a greater similarity to ulcerative colitis and showed stronger associations with *Bacteroides vulgatus* and neutrophil activity, whereas ileal CD was more closely linked to bile acid metabolism and demonstrated marked alterations in *Faecalibacterium prausnitzii*. These findings underscore how both the model choice and feature modality substantially influence predictive performance.⁶²

Besides single-algorithm approaches, ensemble methods are also advantageous. Ha et al trained eight ML algorithms, including RF, deep neural networks (DNNs), logistic regression, k-nearest neighbors, decision trees, gradient boosting, and SVMs, on multi-cohort pediatric IBD data. An ensemble model combining three top-performing classifiers (DNN, logistic regression, and SVMs) achieved the highest accuracy in predicting future remission, outperforming any single algorithm.⁵² The characteristics of each ML/DL algorithm are summarized in Table 1.

Table I Comparative Characteristics and Performance Trends of ML Algorithms in IBD Microbiome Research

Algorithm	Key Strengths	Key Limitations
Random Forest (RF) ^{48,57}	• Non-parametric ensemble that reduces variance via bootstrap aggregation (bagging). • Built-in out-of-bag (OOB) validation provides internal error estimates without a separate test set. • Generates feature importance metrics. • Requires minimal data processing; robust to missing values.	• Limited global interpretability (black-box ensemble). • Correlated predictors may bias feature importance. • Class imbalance or over-tuning (tree depth, number) can cause instability.
Support Vector Machine (SVM) ^{63,64}	• Constructs maximal-margin hyperplanes • Effective in high-dimensional, low-sample data. • Kernel trick enables flexible non-linear classification.	• Sensitive to kernel and hyperparameter (C, γ) selection • Requires tuning for different omics scales. • Poor scalability for large multi-omics matrices.
Boosting (eg, XGBoost, AdaBoost, LightGBM) ^{18,65}	• Sequentially builds additive weak learners to minimize residual error, improving bias–variance trade-off. • Incorporates regularization and sparsity awareness (eg, XGBoost) to prevent overfitting.	• Prone to overfitting on small or noisy microbiome datasets if not properly regularized. • Large hyper-parameter space increases computational demand. • Difficult to parallelize due to sequential dependence (less of an issue in LightGBM).
Deep Learning (DNN, CNN, Autoencoder) ^{66,67}	• Learns hierarchical latent features automatically from raw compositional or multi-omics data. • Captures complex non-linear dependencies and cross-modal correlations. • Excellent performance on unstructured data (image, text, speech).	• Requires large, well-labeled datasets • Sensitive to overfitting. • Low interpretability (“black-box”) and difficult hyper-parameter optimization. • Requires significant computational resources (GPU/TPU) and training time.
Ensemble (Stacking / Bagging / Voting / Hybrid) ^{48,52,68}	• Aggregates heterogeneous models (eg, RF, SVM, DL) to exploit complementary strengths and improve stability. • Reduces prediction variance and enhances cross-cohort robustness.	• Increased model complexity makes calibration and interpretation difficult. • Risk of overfitting when base learners are correlated or data are limited.

Challenges and Limitations of ML in IBD Microbiome Research

However, despite these advances, several challenges remain unresolved. Overfitting is a recurrent issue due to the imbalance between high-dimensional microbial features and relatively small sample sizes. A recent review by Dudek et al emphasized the risks of data leakage and overfitting in microbiome ML workflows, calling for more rigorous cross-validation and model-locking strategies.⁶⁹ In addition to overfitting, performance metrics such as AUCs or accuracy often rely solely on internal cross-validation and may overestimate true generalizability. Robust evaluation requires nested cross-validation, bootstrapping, and, ideally, independent external validation to ensure fair model assessment and prevent information leakage.^{48,70} As demonstrated by Kubinski et al, model performance varies markedly across preprocessing choices and cohorts, underscoring the difficulty of achieving cross-cohort reproducibility.⁷¹ Moreover, upstream sources of technical noise—including differences in normalization procedures, zero-inflated feature distributions, and the inherent sparsity and compositionality of microbial data—can substantially distort model behavior and feature importance rankings, further complicating reproducibility and external validation.^{71–73} Furthermore, as highlighted by Papoutsoglou et al, the lack of standardized preprocessing pipelines and benchmarking datasets undermines reproducibility and complicates comparisons across studies.⁴⁸ Collectively, these findings demonstrate that while ML has brought substantial progress beyond traditional statistical approaches in IBD microbiome research, persistent barriers, including overfitting, limited generalizability, and lack of methodological standardization, must be addressed. Rigorous external validation, harmonized data pipelines, and the integration of interpretable ML frameworks are essential to ensure that microbiome-based models can transition into reliable clinical applications.⁷⁴

Emerging Approaches Beyond Traditional ML

While traditional ML approaches such as RF and SVMs have shown strong utility in IBD microbiome research, recent years have seen the rise of deep learning (DL) and next-generation AI frameworks that aim to capture even more complex and nonlinear relationships within multi-omics data.^{66,75} Unlike traditional methods that often rely on predefined or engineered features, DL methods use representation learning to automatically extract latent patterns from raw or minimally processed inputs.^{75,76} Oh and Zhang reported DeepMicro as a notable early framework that employed autoencoders to compress high-dimensional microbiome profiles into compact latent representations, thereby enhancing the disease classification performance. Building on this foundation, subsequent DL applications have explored a range of architectures: autoencoders for dimensionality reduction and feature learning,⁷⁵ and convolutional neural networks (CNNs) to identify structured patterns in microbial abundance matrices.⁷⁷ More advanced models, such as the Multimodal Variational Information Bottleneck (MVIB), further integrate multi-omics layers and clinical metadata, highlighting how DL can support more accurate and clinically relevant predictions.⁷⁸ Collectively, these approaches emphasize the potential of DL to surpass conventional ML, offering scalable and flexible tools for predictive microbiome research.

Recently, attention has shifted toward the development of foundation models that leverage large-scale pretraining on biological sequences to generate generalized representations. The Microbial General Model (MGM) was one of the first attempts to construct a microbiome-focused foundation model trained on more than 260,000 microbiome samples using transformer-based pretraining.⁷⁹ Although still in its early stages, the MGM demonstrates how models tailored to microbial communities can eventually classify community types, identify keystone taxa, and capture longitudinal dynamics, thereby providing a scalable framework for studying microbiome-driven processes in IBD. In parallel, host-centric genome foundation models, such as AlphaGenome⁸⁰ and Nucleotide Transformer,⁸¹ have been developed to annotate genetic variants and predict their regulatory effects in human genomics. Although not originally designed for microbial applications, these host-oriented models provide transferable embeddings that may help bridge host genetic information with microbial-community dynamics. Together, microbiome- and host-focused foundation models highlight a forward-looking trajectory in which unified representational spaces may enable multi-omics integration and deepen our understanding of host–microbiome interactions in IBD.

In parallel, recent multi-omics studies have begun to adopt time-series and causal inference models that capture longitudinal microbiome–host interactions and reveal potential drivers of disease dynamics.^{82,83} Dynamic Bayesian networks, Granger-causality analysis, and structural causal modeling frameworks allow researchers to infer temporal dependencies and directionality in microbial and metabolic shifts, providing insights that static cross-sectional analyses cannot capture.⁸⁴ These temporal and causal modeling approaches are crucial for understanding how microbial community states evolve during inflammation, remission, and therapeutic intervention in IBD. In addition, multi-omics data-fusion frameworks have rapidly advanced predictive microbiome analytics by integrating heterogeneous data layers within shared latent spaces. Representative examples include Multi-Omics Factor Analysis (MOFA+), which applies probabilistic factor modeling to uncover coordinated variance across omics layers;⁸⁵ DIABLO, a multivariate approach for discriminative integration of multi-omic modalities;⁸⁶ and deep representation learning frameworks, such as variational autoencoder–based or multimodal transformer architectures, which embed genomic, transcriptomic, and metabolomic profiles into unified representations for improved disease prediction and mechanistic interpretation.⁷⁵ More recently, the MintTea (Multi-omics Integration through Temporal Embedding and Attention) framework has extended this concept by integrating longitudinal multi-omics data through attention-based deep learning, enabling the identification of dynamic, disease-associated molecular modules across time and modalities.⁸⁷ These integrative frameworks exemplify the shift from isolated single-omic analysis toward data-fusion models that bridge microbial, metabolic, and host dimensions of IBD.

As these models grow in complexity, concerns regarding their interpretability have become more pressing. To ensure clinical relevance and acceptance, the integration of explainable AI (XAI)⁸⁸ methods, such as SHapley Additive exPlanation (SHAP) values, attention weight visualization, and concept bottleneck models, has been proposed.⁸⁹ For instance, Novielli et al applied SHAP values to microbiome classification tasks and showed that taxa such as *Fusobacterium* and *Parvimonas* were key drivers of colorectal cancer predictions, thereby providing transparent links between features and model outputs.⁹⁰ Similarly, Onwuka et al demonstrated that the SHAP-based prioritization of fecal and plasma metabolites could highlight IBD-associated signatures, reinforcing the role of XAI in biomarker discovery and clinician application.⁹¹ In line with these advances, a recent review by Kim et al emphasized that the field is shifting

from descriptive profiling of dysbiosis toward more predictive and analytic (including prescriptive) applications, while highlighting ongoing challenges such as biological and technical heterogeneity, limited generalizability across cohorts, and the need for methodological standardization and external validation.⁹²

Clinical Translation and Challenges

The application of ML and DL to the IBD microbiome data has generated promising diagnostic models and candidate biomarker panels.^{93,94} However, the path from computational discovery to clinical implementation remains complex and requires careful consideration of several limitations.^{95,96} While traditional statistical analyses and ML/DL-based approaches have successfully identified bacterial taxa associated with disease states, causal or mechanistic links between these microbes and host physiology have been established only for a limited subset. Most computational analyses cannot by themselves determine whether these microbial signatures drive, result from, or merely correlate with inflammation. For example, several genera including *Faecalibacterium* frequently identified by differential abundant taxa analysis are known to produce short-chain fatty acids that modulate intestinal immune responses, and their depletion may contribute to mucosal inflammation and barrier dysfunction in IBD.^{97,98} Conversely, enrichment of *Escherichia-Shigella* or *Enterococcus* may reflect pro-inflammatory activity through lipopolysaccharide-mediated TLR signaling.^{99,100} To validate and interpret these associations, further experimental confirmation through in vivo or ex vivo studies is required, as well as the use of next-generation computational frameworks—such as foundation-model-based simulation or causal network modeling—to infer and test host–microbe interaction dynamics.¹⁰¹

Standardization and Reproducibility of Microbiome Assays

A key challenge is the standardization of microbiome assays. Pre-analytical variables such as sample types (feces or biopsies), stool collection methods, preservation techniques, sequencing platforms, and library preparation introduce substantial variability that directly affects downstream feature profiles.

In particular, differences in sequencing chemistry, read length, and bioinformatic preprocessing pipelines can introduce systematic biases in taxonomic resolution and relative abundance estimation.^{102,103} Batch effects arising from DNA extraction kits, primer selection, and library preparation protocols have been shown to exceed inter-individual biological variability in some cases, thereby confounding disease-associated microbial signatures across studies.¹⁰⁴

Together, these technical factors interact with sampling- and host-related sources of variability, underscoring the importance of methodological harmonization. Kim et al compared the gut microbiome profiles obtained from the stool, luminal contents, and mucosal biopsies of patients with UC and healthy controls.¹⁰⁵ They found significant differences in both alpha and beta diversities across sample types, with biopsies yielding higher numbers of observed OTUs than stool or lavage samples. Community structures showed a correlation between stool and luminal contents, whereas biopsy samples did not correlate with either. Importantly, UC-control differences were evident only in stool samples, whereas lavage and biopsy samples did not show significant separation, partly due to the limited sample size. Classification analysis achieved AUCs of 0.85 for stool and 0.81 for lavage, underscoring that the sampling strategy critically influences microbiome readouts and diagnostic performance. Kruger et al systematically assessed intra- and inter-individual variability in gut health markers in healthy adults using an optimized fecal sampling and processing workflow.¹⁰⁶ They showed that the variability differed by marker, with some exhibiting substantial within-person variations. Importantly, they demonstrated that optimized pre-processing methods, such as mill homogenization, could significantly reduce technical variability, highlighting that methodological choices can obscure biological signals if not standardized. This underscores the need for assay harmonization to ensure reliable biomarker discovery and provide a robust foundation for ML applications in clinical contexts. Similarly, Clooney et al performed a large longitudinal multicenter study across two continents in patients with IBD and controls, revealing that the geographic site and temporal factors explained a significant share of the microbiome variance, often rivaling or exceeding disease effects.¹⁰⁷ They further observed that temporal stability was reduced in IBD, and that model performance improved when consecutive time points were considered. These findings illustrate how cross-site and temporal heterogeneity can mask disease-associated signatures, and emphasize the importance of standardized methodologies and rigorous external validation in the development of ML-based diagnostics and prognostics for IBD. Unless these sources of heterogeneity are rigorously controlled, models trained in one context are unlikely to be generalized to other clinical settings.¹⁰⁸ This challenge

becomes even more pronounced when integrating multi-omics layers, such as metabolomic, transcriptomic, and proteomic data, which adds further complexity to model development and validation.

Model Interpretability and Explainable AI

Another major concern was the interpretability of the model. Algorithms such as RF, boosting algorithms, and neural networks often achieve high predictive performance but can act as “black boxes”.⁹⁶ In clinical practice, the opacity of AI models can undermine the trust of clinicians and complicate regulatory approval. A systematic review by Rosenbacke et al reported that explainable AI can either enhance or diminish trust depending on how explanations are presented.¹⁰⁹ Explanations that are clear, concise, and consistent with clinical reasoning tend to increase clinicians’ confidence in AI systems. Conversely, explanations that are overly technical, complex, inconsistent with clinical logic, or misleading can reduce it. Thus, providing transparent and clinically relevant explanations is essential for successfully integrating AI into healthcare and securing the trust of medical professionals. Rynazal et al applied XAI methods (notably SHAP for local explanations) to colorectal cancer (CRC) classification using gut microbiome data.¹¹⁰ They demonstrated that using local (patient-specific) feature attributes revealed which microbial taxa were most influential for individual predictions and enabled the stratification of patients with CRC into subgroups, thereby improving interpretability and potentially pointing toward mechanistic hypotheses. Yu et al developed a 10-species microbial signature for IBD using an XGBoost-based model (termed XGB-IBD10) and performed external validation.¹¹¹ Importantly, they applied SHAP analysis to explain model predictions, showing how individual taxa influenced the classification decisions. External validation confirmed the robustness of the identified microbial panel, demonstrating that explainability could enhance trust in model outputs and provide mechanistic insights into disease-associated microbes.

Taken together, these studies illustrate that XAI approaches, including feature importance ranking, SHAP values, and attention-based visualization, are critical not only for identifying which microbial taxa, metabolic pathways, or host features drive predictions but also for facilitating clinical adoption and guiding biological hypothesis generation. This process can provide a reliable output for clinical deployment.

Demonstrating Clinical Utility and Path to Adoption

Finally, the demonstration of its clinical utility remains a missing link. Current studies primarily focus on metrics such as the AUC, sensitivity, and specificity. Although these benchmarks are essential for the initial validation, they fail to capture whether such models can improve patient care. To achieve regulatory approval and widespread clinical uptake, predictive frameworks must demonstrate tangible benefits, such as reducing unnecessary colonoscopies, expediting the time to therapy initiation, and lowering healthcare costs. Ultimately, prospective validation and interventional studies are required to bridge the gap between computational promise and real-world clinical impact.

Recent longitudinal research has underscored this need. Al Radi et al investigated the predictive value of gut microbiome signatures for therapy intensification in IBD using a 10-year follow-up cohort.¹¹² Their findings demonstrated that baseline microbial features could predict the need for treatment escalation, highlighting the importance of long-term monitoring and the value of microbiome-informed risk stratification for guiding clinical decision making. This study illustrated how longitudinal datasets provide a rigorous framework for evaluating the durability and clinical relevance of microbiome-derived predictors. In parallel, Massaro et al presented the design of the OPTIMIST study, a prospective, longitudinal, observational pilot study conducted in Canada to investigate gut microbiome predictors of advanced therapy response in CD.¹¹³ The study enrolled patients initiating biological treatment, as well as non-IBD controls, to link baseline microbiome composition and temporal dynamics to therapeutic outcomes. By prospectively monitoring patients over time, the OPTIMIST study established a framework for evaluating microbiome-based predictors in real-world settings and supporting their eventual clinical translation.

In summary, ML and DL hold considerable promise for translating microbiome research into diagnostic and prognostic tools for IBD. However, clinical translation will require overcoming persistent challenges related to reproducibility across cohorts, generalizability to diverse patient populations, interpretability of results, and addressing data-sharing constraints. Progress in explainable AI, development of standardized pipelines, and adherence to rigorous reporting and evaluation guidelines will be crucial in bridging the gap between proof-of-concept studies and real-world patient care.

Conclusion and Future Perspectives

ML and DL have transformed IBD microbiome research from descriptive community profiling to predictive modeling. Alongside DA analyses, these approaches have enabled the classification of disease versus health, the differentiation of CD from UC, and the identification of microbial and metabolic biomarkers with diagnostic and prognostic significance (Figure 1). Studies employing RF, boosting frameworks, and ensemble models have demonstrated that microbiome-informed classifiers can approach or even surpass traditional biomarkers, and explainable AI methods are beginning to provide clearer insights into the microbial and functional drivers of diseases.

Despite this progress, significant barriers remain before clinical adoption can be achieved. Variability in stool collection, sequencing platforms, and preprocessing pipelines continues to hinder reproducibility, and external validation often reveals a decline in model performance when applied across independent cohorts. Dependence on complex “black-box” algorithms also makes it more difficult for clinicians to trust and regulators to approve. Furthermore, most studies are retrospective, with few prospective or interventional studies directly evaluating whether microbiome-based ML models can improve patient outcomes. These issues emphasize the need to connect computational advances with the development of practical tools for precision medicine.

Looking ahead, progress will be fastest if efforts focus on: (1) rigorous validation, prioritizing pre-registered pipelines, locked models, and multi-site external testing; (2) strong benchmarking, with standardized preprocessing, shared baselines, and challenge datasets covering stool and mucosal samples, platforms, and regions; (3) biologically based interpretability, pairing explanation methods with wet-lab or orthogonal assays to verify mechanistic plausibility; (4) longitudinal and multimodal integration, combining metagenomics, metabolomics, proteomics, and host features to model disease trajectories rather than single time points; and (5) next-generation AI, where emerging genome foundation models—both microbiome-focused and host-centric—are developed under strict evaluation and compared directly to robust classical baselines. These steps will help ML move from promising prototypes to dependable clinical tools that guide diagnosis, monitoring, and treatment decisions for IBD.

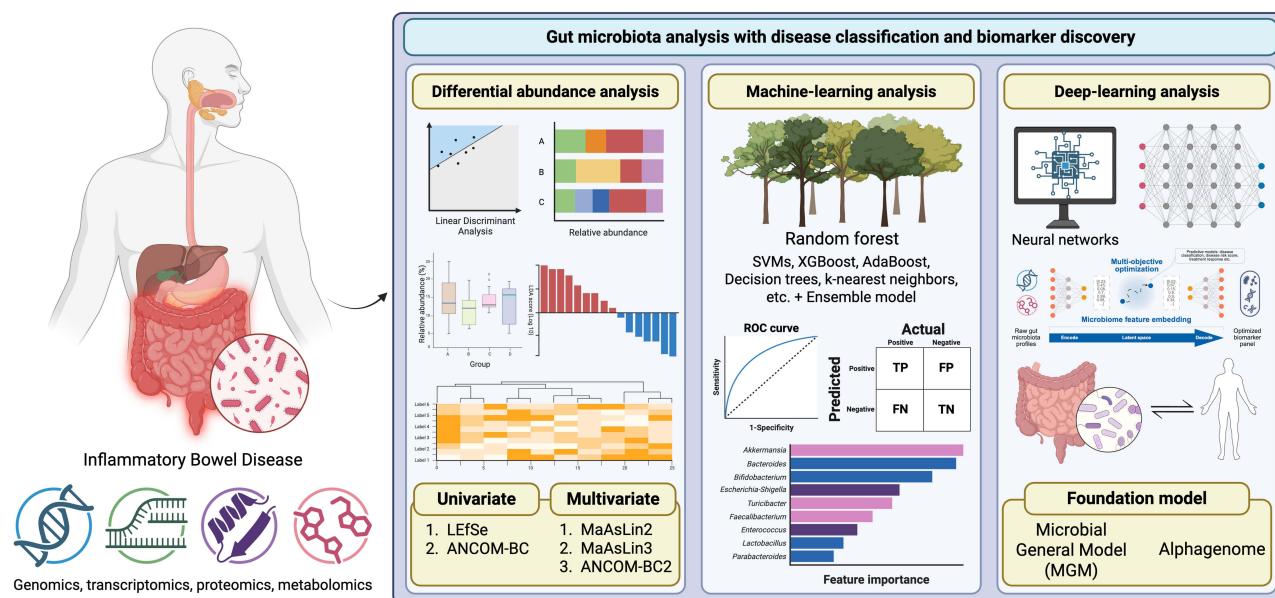


Figure 1 Evolution of analytical approaches for characterizing gut microbiota alterations in inflammatory bowel disease (IBD). An overview of analytical strategies for IBD illustrates the progression from traditional statistics to next-generation artificial intelligence (AI) frameworks. Early differential abundance (DA) analyses using 16S rRNA or metagenomic data identified disease-associated taxa; however, these were often limited by univariate contrasts and reproducibility issues. With the increasing data complexity of multi-omics approaches, machine learning (ML) approaches such as random forests, support vector machines (SVMs), and boosting algorithms have enabled robust disease classification, biomarker panel identification, and performance evaluation through receiver operating characteristic (ROC) curves, cross-validation, and feature importance scores. Recently, deep learning (DL) frameworks have been applied to capture nonlinear relationships in high-dimensional multi-omics datasets, providing latent feature embeddings and integrative predictive models. Looking ahead, foundation models trained on large-scale host and microbial genomic resources (eg. Microbial General Model for microbiome and AlphaGenome for host genetics) are emerging as transferable backbones that can be fine-tuned for diverse downstream tasks, from diagnosis and subtype prediction to therapy response modeling and prognosis.

Data Sharing Statement

Data sharing not applicable – no new data generated.

Acknowledgments

The figure was generated using BioRender by June-Young Lee (2025) <https://BioRender.com/je7gx57>. We would like to thank Editage for the English language editing.

Author Contributions

- June-Young Lee (J.Y. Lee): Conceptualization; Methodology; Formal analysis; Software; Data curation; Investigation; Resources; Visualization; Funding acquisition; Writing – original draft; Writing – review & editing.
- Dong Hyun Kim (D.H. Kim): Formal analysis; Software; Validation; Data curation; Visualization; Writing – review & editing.
- Jee-Won Choi (J.W. Choi): Data curation; Investigation; Visualization; Writing – review & editing.
- Minh Shong (M. Shong): Resources; Investigation; Writing – review & editing.
- Chang Kyun Lee (C.K. Lee): Conceptualization; Methodology; Supervision; Project administration; Funding acquisition; Writing – review & editing.

All authors contributed substantially to the work, critically revised the manuscript, and agree to be accountable for all aspects of the work. All authors gave final approval of the version to be published. June-Young Lee and Dong Hyun Kim contributed equally to this work and share first authorship.

Funding

This research was supported by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (grant number: RS-2023-KH135855) and the InnoCORE program of the Ministry of Science and ICT, Republic of Korea (grant number: N10250153).

Disclosure

The authors declare that they have no competing interests.

References

1. Ni J, Wu GD, Albenberg L, Tomov VT. Gut microbiota and IBD: causation or correlation? *Nat Rev Gastroenterol Hepatol*. 2017;14(10):573–584. doi:10.1038/nrgastro.2017.88
2. Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology*. 2014;146(6):1489–1499. doi:10.1053/j.gastro.2014.02.009
3. Shan Y, Lee M, Chang EB. The gut microbiome and inflammatory bowel diseases. *Annual Review of Medicine*. 2022;73:455–468. doi:10.1146/annurev-med-042320-021020
4. Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature*. 2019;569(7758):655–662. doi:10.1038/s41586-019-1237-9
5. Morgan XC, Tickle TL, Sokol H, et al. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol*. 2012;13(9):R79. doi:10.1186/gb-2012-13-9-r79
6. Ning L, Zhou Y-L, Sun H, et al. Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nat Commun*. 2023;14(1):7135. doi:10.1038/s41467-023-42788-0
7. Zhang Y, Thomas JP, Korcsmaros T, Gul L. Integrating multi-omics to unravel host-microbiome interactions in inflammatory bowel disease. *Cell Rep Med*. 2024;5(9):101738. doi:10.1016/j.xcrm.2024.101738
8. Hu S, Bourgonje AR, Gacesa R, et al. Mucosal host-microbe interactions associate with clinical phenotypes in inflammatory bowel disease. *Nat Commun*. 2024;15(1):1470. doi:10.1038/s41467-024-45855-2
9. Jayasooriya N, Baillie S, Blackwell J, et al. Systematic review with meta-analysis: time to diagnosis and the impact of delayed diagnosis on clinical outcomes in inflammatory bowel disease. *Aliment. Pharmacol. Ther*. 2023;57(6):635–652. doi:10.1111/apt.17370
10. Zhang SY, Lin Y. Addressing diagnostic delays in inflammatory bowel diseases in Germany. *World J Gastroenterol*. 2024;30(41):4484–4489. doi:10.3748/wjg.v30.i41.4484
11. Henriksen M, Jahnsen J, Lygren I, et al. C-reactive protein: a predictive factor and marker of inflammation in inflammatory bowel disease. Results from a prospective population-based study. *Gut*. 2008;57(11):1518–1523. doi:10.1136/gut.2007.146357
12. Walsham NE, Sherwood RA. Fecal calprotectin in inflammatory bowel disease. *Clin Exp Gastroenterol*. 2016;9:21–29. doi:10.2147/ceg.S51902

13. Clough J, Colwill M, Poullis A, Pollok R, Patel K, Honap S. Biomarkers in inflammatory bowel disease: a practical guide. *Ther Adv Gastroenterol*. 2024;17:17562848241251600. doi:10.1177/17562848241251600
14. Verstockt B, Noor NM, Marigorta UM, Pavlidis P, Deepak P, Ungaro RC. Results of the seventh scientific workshop of ECCO: precision medicine in IBD-disease outcome and response to therapy. *J Crohns Colitis*. 2021;15(9):1431–1442. doi:10.1093/ecco-jcc/jjab050
15. Chen L, Zhang C, Niu R, et al. Multi-omics biomarkers for predicting efficacy of biologic and small-molecule therapies in adults with inflammatory bowel disease: a systematic review. *United Eur Gastroenterol J*. 2025;13(4):517–530. doi:10.1002/ueg2.12720
16. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput Biol*. 2021;17(11):e1009442. doi:10.1371/journal.pcbi.1009442
17. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60. doi:10.1186/gb-2011-12-6-r60
18. Hernández Medina R, Kutuzova S, Nielsen KN, et al. Machine learning and deep learning applications in microbiome research. *ISME Commun*. 2022;2(1):98. doi:10.1038/s43705-022-00182-9
19. Nearing JT, Douglas GM, Hayes MG, et al. Microbiome differential abundance methods produce different results across 38 datasets. *Nat Commun*. 2022;13(1):342. doi:10.1038/s41467-022-28034-z
20. Halfvarson J, Brislawn CJ, Lamendella R, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol*. 2017;2(5):17004. doi:10.1038/nmicrobiol.2017.4
21. Franzosa EA, Sirota-Madi A, Avila-Pacheco J, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol*. 2019;4(2):293–305. doi:10.1038/s41564-018-0306-4
22. Tarca AL, Carey VJ, X-w C, Romero R, Drăghici S. Machine Learning and Its Applications to Biology. *PLoS Comput Biol*. 2007;3(6):e116. doi:10.1371/journal.pcbi.0030116
23. Marcos-Zambrano LJ, Karadzovic-Hadziabdic K, Loncar Turukalo T, et al. Applications of machine learning in human microbiome studies: a review on feature selection, biomarker identification, disease prediction and treatment. Review. *Front Microbiol*. 2021;Volume 12. doi:10.3389/fmicb.2021.634511
24. Zheng J, Sun Q, Zhang M, et al. Noninvasive, microbiome-based diagnosis of inflammatory bowel disease. *Nat Med*. 2024;30(12):3555–3567. doi:10.1038/s41591-024-03280-4
25. Khachatryan L, Xiang Y, Ivanov A, et al. Results and lessons learned from the sbv IMPROVER metagenomics diagnostics for inflammatory bowel disease challenge. *Sci Rep*. 2023;13(1):6303. doi:10.1038/s41598-023-33050-0
26. Caminero A, Tropini C, Valles-Colomer M, et al. Credible inferences in microbiome research: ensuring rigour, reproducibility and relevance in the era of AI. *Nat Rev Gastroenterol Hepatol*. 2025;22(11):788–803. doi:10.1038/s41575-025-01100-9
27. Gilbert JA, Azad MB, Bäckhed F, et al. Clinical translation of microbiome research. *Nat Med*. 2025;31(4):1099–1113. doi:10.1038/s41591-025-03615-9
28. Ma X, Lu X, Zhang W, et al. Gut microbiota in the early stage of Crohn's disease has unique characteristics. *Gut Pathog*. 2022;14(1):46. doi:10.1186/s13099-022-00521-0
29. Nickols WA, Kuntz T, Shen J, et al. MaAsLin 3: refining and extending generalized multivariable linear models for meta-omic association discovery. *bioRxiv*. 2024;2024.12.13.628459. doi:10.1101/2024.12.13.628459
30. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. *Nat Commun*. 2020;11(1):3514. doi:10.1038/s41467-020-17041-7
31. Lin H, Peddada SD. Multigroup analysis of compositions of microbiomes with covariate adjustments and repeated measures. *Nat Method*. 2024;21(1):83–91. doi:10.1038/s41592-023-02092-7
32. Chen W, Li Y, Wang W, et al. Enhanced microbiota profiling in patients with quiescent Crohn's disease through comparison with paired healthy first-degree relatives. *Cell Rep Med*. 2024;5(7):101624. doi:10.1016/j.xcrm.2024.101624
33. Rosso AD, Aguilera P, Quesada S, et al. Comprehensive phenotyping in inflammatory bowel disease: search for biomarker algorithms in the transkingdom interactions context. *Microorganisms*. 2022;10(11):2190. doi:10.3390/microorganisms10112190
34. Gloor GB. ALDEX 2: ANOVA-Like Differential Expression tool for compositional data. 2016.
35. Zhou H, He K, Chen J, Zhang X. LinDA: linear models for differential abundance analysis of microbiome compositional data. *Genome Biol*. 2022;23(1):95. doi:10.1186/s13059-022-02655-5
36. Vich Vila A, Imhann F, Collij V, et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci Transl Med*. 2018;10(472). doi:10.1126/scitranslmed.aap8914
37. Akiyama S, Nishijima S, Kojima Y, et al. Multi-biome analysis identifies distinct gut microbial signatures and their crosstalk in ulcerative colitis and Crohn's disease. *Nat Commun*. 2024;15(1):10291. doi:10.1038/s41467-024-54797-8
38. Tun HM, Peng Y, Massimino L, et al. Gut virome in inflammatory bowel disease and beyond. *Gut*. 2024;73(2):350–360. doi:10.1136/gutjnl-2023-330001
39. Norman Jason M, Handley Scott A, Baldrige Megan T, et al. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell*. 2015;160(3):447–460. doi:10.1016/j.cell.2015.01.002
40. Hsu C, Ghannoum M, Cominelli F, Martino LD. Mycobiome and inflammatory bowel disease: role in disease pathogenesis, current approaches and novel nutritional-based therapies. *Inflamm Bowel Dis*. 2023;29(3):470–479. doi:10.1093/ibd/izac156
41. Zhang X, Deeke SA, Ning Z, et al. Metaproteomics reveals associations between microbiome and intestinal extracellular vesicle proteins in pediatric inflammatory bowel disease. *Nat Commun*. 2018;9(1):2873. doi:10.1038/s41467-018-05357-4
42. Li M, Yang L, Mu C, et al. Gut microbial metabolome in inflammatory bowel disease: from association to therapeutic perspectives. *Comput Struct Biotechnol J*. 2022;20:2402–2414. doi:10.1016/j.csbj.2022.03.038
43. Liu L, Wang G, Safo SE. Multivariate sparse functional linear discriminant analysis: an application to inflammatory bowel disease classification. *arXiv preprint, arXiv*. 2025.
44. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. Mini review. *Front Microbiol*. 2017;8. doi:10.3389/fmicb.2017.02224
45. Silverman JD, Washburne AD, Mukherjee S, David LA. A phylogenetic transform enhances analysis of compositional microbiota data. *eLife*. 2017;6:e21887. doi:10.7554/eLife.21887
46. Hodgkiss R, Acharjee A. Unravelling metabolite-microbiome interactions in inflammatory bowel disease through AI and interaction-based modelling. *Biochimica et Biophysica Acta*. 2025;1871(3):167618. doi:10.1016/j.bbdis.2024.167618
47. Manandhar I, Alimadadi A, Aryal S, Munroe PB, Joe B, Cheng X. Gut microbiome-based supervised machine learning for clinical diagnosis of inflammatory bowel diseases. *Am J Physiol Gastrointest Liver Physiol*. 2021;320(3):G328–g337. doi:10.1152/ajpgi.00360.2020

48. Papoutsoglou G, Tarazona S, Lopes MB, et al. Machine learning approaches in microbiome research: challenges and best practices. Review. *Front Microbiol.* **2023**;14. doi:10.3389/fmicb.2023.1261889
49. Lee S, Lee I. Comprehensive assessment of machine learning methods for diagnosing gastrointestinal diseases through whole metagenome sequencing data. *Gut Microbes.* **2024**;16(1):2375679. doi:10.1080/19490976.2024.2375679
50. Liñares-Blanco J, Fernandez-Lozano C, Seoane JA, López-Campos G. Machine learning based microbiome signature to predict inflammatory bowel disease subtypes. original research. *Front Microbiol.* **2022**;13. doi:10.3389/fmicb.2022.872671
51. Kim H, Na JE, Kim S, et al. A machine learning-based diagnostic model for crohn's disease and ulcerative colitis utilizing fecal microbiome analysis. *Microorganisms.* **2023**;12(1):36. doi:10.3390/microorganisms12010036
52. Ha SM, Lee K, Kim GH, Hurych J, Cinek O, Shim JO. Gut-microbiota-based ensemble model predicts prognosis of pediatric inflammatory bowel disease. *Iscience.* **2024**;27(12):111442. doi:10.1016/j.isci.2024.111442
53. Jacobs JP, Sauk JS, Ahdoot AI, et al. Microbial and metabolite signatures of stress reactivity in ulcerative colitis patients in clinical remission predict clinical flare risk. *Inflamm Bowel Dis.* **2023**;30(3):336–346. doi:10.1093/ibd/izad185
54. Kim HS, Oh SJ, Kim BK, et al. Dysbiotic signatures and diagnostic potential of gut microbial markers for inflammatory bowel disease in Korean population. *Sci Rep.* **2024**;14(1):23701. doi:10.1038/s41598-024-74002-6
55. Lee Y, Cappellato M, Di Camillo B. Machine learning-based feature selection to search stable microbial biomarkers: application to inflammatory bowel disease. *Gigascience.* **2022**;12. doi:10.1093/gigascience/giad083
56. Abbas M, Matta J, Le T, et al. Biomarker discovery in inflammatory bowel diseases using network-based feature selection. *PLoS One.* **2019**;14(11):e0225382. doi:10.1371/journal.pone.0225382
57. Breiman L. Random Forests. *Mach Learn.* **2001**;45(1):5–32. doi:10.1023/A:1010933404324
58. Statnikov A, Henaff M, Narendra V, et al. A comprehensive evaluation of multicategory classification methods for microbiomic data. *Microbiome.* **2013**;1(1):11. doi:10.1186/2049-2618-1-11
59. Knights D, Costello EK, Knight R. Supervised classification of human microbiota. *FEMS Microbiol Rev.* **2011**;35(2):343–359. doi:10.1111/j.1574-6976.2010.00251.x
60. Li L, Deng X, Wang S, Huang T. Integrating traditional omics and machine learning approaches to identify microbial biomarkers and therapeutic targets in pediatric inflammatory bowel disease. *Front Pharmacol.* **2025**;16:1545392. doi:10.3389/fphar.2025.1545392
61. Li L, Liu Z, Qin J, et al. Constructing inflammatory bowel disease diagnostic models based on k-mer and machine learning. *Front Microbiol.* **2025**;16:1578005. doi:10.3389/fmicb.2025.1578005
62. Gonzalez CG, Mills RH, Zhu Q, et al. Location-specific signatures of Crohn's disease at a multi-omics scale. *Microbiome.* **2022**;10(1):133. doi:10.1186/s40168-022-01331-x
63. Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* **1995**;20(3):273–297. doi:10.1007/BF00994018
64. Jiang Y, Luo J, Huang D, Liu Y, Li DD. Machine learning advances in microbiology: a review of methods and applications. *Front Microbiol.* **2022**;13:925454. doi:10.3389/fmicb.2022.925454
65. Jerome HF. Greedy function approximation: a gradient boosting machine. *Ann Stat.* **2001**;29(5):1189–1232. doi:10.1214/aos/1013203451
66. Przyms P, Rykaczewski K, Martin-Segura A, et al. Deep learning in microbiome analysis: a comprehensive review of neural network models. *Rev Front Microbiol.* **2025**;15. doi:10.3389/fmicb.2024.1516667
67. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* **2015**;521(7553):436–444. doi:10.1038/nature14539
68. Polikar R. Ensemble based systems in decision making. *IEEE Circuits Syst Mag.* **2006**;6(3):21–45. doi:10.1109/MCAS.2006.1688199
69. Dudek NK, Chakhvadze M, Kobakhidze S, Kantidze O, Gankin Y. Supervised machine learning for microbiomics: bridging the gap between current and best practices. *Mach Learn Appl.* **2024**;18:100607. doi:10.1016/j.mlwa.2024.100607
70. Topçuoğlu BD, Lesniak NA, MTt R, Wiens J, Schloss PD. A framework for effective application of machine learning to microbiome-based classification problems. *mBio.* **2020**;11(3). doi:10.1128/mBio.00434-20
71. Kubinski R, Djamen-Kepaou JY, Zhanabaev T, et al. Benchmark of data processing methods and machine learning models for gut microbiome-based diagnosis of inflammatory bowel disease. *Front Genet.* **2022**;13:784397. doi:10.3389/fgene.2022.784397
72. Lin H, Peddada SD. Analysis of microbial compositions: a review of normalization and differential abundance analysis. *Npj Biofilms Microbi.* **2020**;6(1):60. doi:10.1038/s41522-020-00160-w
73. Pan AY. Statistical analysis of microbiome data: the challenge of sparsity. *Curr. Opin. Endocr. Metab. Res.* **2021**;19:35–40. doi:10.1016/j.coemr.2021.05.005
74. Zand A, Stokes Z, Sharma A, van Deen WK, Hommes D. Artificial intelligence for Inflammatory Bowel Diseases (IBD); accurately predicting adverse outcomes using machine learning. *Dig Dis Sci.* **2022**;67(10):4874–4885. doi:10.1007/s10620-022-07506-8
75. Oh M, Zhang L. DeepMicro: deep representation learning for disease prediction based on microbiome data. *Sci Rep.* **2020**;10(1):6026. doi:10.1038/s41598-020-63159-5
76. Jain S, Safo SE. DeepIDA-GRU: a deep learning pipeline for integrative discriminant analysis of cross-sectional and longitudinal multiview data with applications to inflammatory bowel disease classification. *Briefings Bioinf* **2024**;25(4). doi:10.1093/bib/bbae339
77. Reiman D, Metwally A, Yang D. Using convolutional neural networks to explore the microbiome. *Annu Int Conf IEEE Eng Med Biol Soc.* **2017**;2017:4269–4272. doi:10.1109/embsc.2017.8037799
78. Grazioli F, Siarheyev R, Alqassem I, Henschel A, Pileggi G, Meiser A. Microbiome-based disease prediction with multimodal variational information bottlenecks. *PLoS Comput Biol.* **2022**;18(4):e1010050. doi:10.1371/journal.pcbi.1010050
79. Zhang H, Zhang Y, Kang Z, Song L, Yang R, Ning K. MGM as a large-scale pretrained foundation model for microbiome analyses in diverse contexts. *bioRxiv.* **2025**. doi:10.1101/2024.12.30.630825
80. Avsec Ž, Latysheva N, Cheng J, et al. AlphaGenome: advancing regulatory variant effect prediction with a unified DNA sequence model. *bioRxiv.* **2025**. doi:10.1101/2025.06.25.661532
81. Dalla-Torre H, Gonzalez L, Mendoza-Revilla J, et al. Nucleotide transformer: building and evaluating robust foundation models for human genomics. *Nat Method.* **2025**;22(2):287–297. doi:10.1038/s41592-024-02523-z
82. Szal M, Stebliankin V, Mathee K, Yoo C, Narasimhan G. Causal effects in microbiomes using interventional calculus. *Sci Rep.* **2021**;11(1):5724. doi:10.1038/s41598-021-84905-3

83. Lyu R, Qu Y, Divaris K, Wu D. Methodological considerations in longitudinal analyses of microbiome data: a comprehensive review. *Genes*. 2024;15(1):51. doi:10.3390/genes15010051
84. Ruiz-Perez D, Lugo-Martinez J, Bourguignon N, et al. Dynamic bayesian networks for integrating multi-omics time series microbiome data. *Msystems*. 2021;6(2). doi:10.1128/mSystems.01105-20
85. Argelaguet R, Velten B, Arnol D, et al. Multi-omics factor analysis-a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol*. 2018;14(6):e8124. doi:10.15252/msb.20178124
86. Singh A, Shannon CP, Gautier B, et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*. 2019;35(17):3055–3062. doi:10.1093/bioinformatics/bty1054
87. Muller E, Shiryani I, Borenstein E. Multi-omic integration of microbiome data for identifying disease-associated modules. *Nat Commun*. 2024;15(1):2621. doi:10.1038/s41467-024-46888-3
88. Sharma C, Sharma S, Sharma K, Sethi GK, Chen H-Y. Exploring explainable AI: a bibliometric analysis. *Discov Appl Sci*. 2024;6(11):615. doi:10.1007/s42452-024-06324-z
89. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. presented at: Proceedings of the 31st International Conference on Neural Information Processing Systems; 2017; Long Beach, California, USA.
90. Novielli P, Romano D, Magarelli M, et al. Explainable artificial intelligence for microbiome data analysis in colorectal cancer biomarker identification. Original Research. *Front Microbiol*. 2024;15. doi:10.3389/fmicb.2024.1348974
91. Onwuka S, Bravo-Merodio L, Gkoutos GV, Acharjee A. Explainable AI-prioritized plasma and fecal metabolites in inflammatory bowel disease and their dietary associations. *Iscience*. 2024;27(7):110298. doi:10.1016/j.isci.2024.110298
92. Kim M, Gim D, Kim S, et al. From dysbiosis to prediction: AI-powered microbiome insights into IBD and CRC. *Gastroenterol Insights*. 2025;16(3):34. doi:10.3390/gastroent16030034
93. Stafford IS, Gosink MM, Mossotto E, Ennis S, Hauben M. A systematic review of artificial intelligence and machine learning applications to inflammatory bowel disease, with practical guidelines for interpretation. *Inflamm Bowel Dis*. 2022;28(10):1573–1583. doi:10.1093/ibd/izac115
94. Tariq R, Afzali A. Artificial intelligence in inflammatory bowel disease: innovations in diagnosis, monitoring, and personalized care. *Therap Adv Gastroenterol*. 2025;18:17562848251357407. doi:10.1177/17562848251357407
95. De Deo D, Dal Buono A, Gabbiadini R, et al. Digital biomarkers and artificial intelligence: a new frontier in personalized management of inflammatory bowel disease. *Rev Front Immunol*. 2025;16. doi:10.3389/fimmu.2025.1637159
96. Tourab H, López Pérez L, Arroyo-Gallego P, et al. The use of machine learning and explainable artificial intelligence in gut microbiome research: a scoping review. 2024.
97. Lenoir M, Martin R, Torres-Maravilla E, et al. Butyrate mediates anti-inflammatory effects of *Faecalibacterium prausnitzii* in intestinal epithelial cells through Dact3. *Gut Microbes*. 2020;12(1):1–16. doi:10.1080/19490976.2020.1826748
98. Zhang Z, Zhang H, Chen T, Shi L, Wang D, Tang D. Regulatory role of short-chain fatty acids in inflammatory bowel disease. *Cell Commun Signal*. 2022;20(1):64. doi:10.1186/s12964-022-00869-5
99. Mirsepasi-Lauridsen HC, Vallance BA, Krogfelt KA, Petersen AM. *Escherichia coli* pathobionts associated with inflammatory bowel disease. *Clin Microbiol Rev*. 2019;32(2). doi:10.1128/cmr.00060-18
100. Giambra V, Pagliari D, Rio P, et al. Gut microbiota, inflammatory bowel disease, and cancer: the role of guardians of innate immunity. *Cells*. 2023;12(22):2654. doi:10.3390/cells12222654
101. Jansma J, El Aidy S. Understanding the host-microbe interactions using metabolic modeling. *Microbiome*. 2021;9(1):16. doi:10.1186/s40168-020-00955-1
102. Nearing JT, Comeau AM, Langille MGI. Identifying biases and their potential solutions in human microbiome studies. *Microbiome*. 2021;9(1):113. doi:10.1186/s40168-021-01059-0
103. Allali I, Arnold JW, Roach J, et al. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol*. 2017;17(1):194. doi:10.1186/s12866-017-1101-8
104. Xu Z, Yeoh YK, Tun HM, et al. Variation in the metagenomic analysis of fecal microbiome composition calls for a standardized operating approach. *Microbiol Spectr*. 2024;12(12):e0151624. doi:10.1128/spectrum.01516-24
105. Kim D, Jung J-Y, Oh H-S, et al. Comparison of sampling methods in assessing the microbiome from patients with ulcerative colitis. *BMC Gastroenterol*. 2021;21(1):396. doi:10.1186/s12876-021-01975-3
106. Kruger K, Myeonghyun Y, van der Wielen N, et al. Evaluation of inter- and intra-variability in gut health markers in healthy adults using an optimised faecal sampling and processing method. *Sci Rep*. 2024;14(1):24580. doi:10.1038/s41598-024-75477-z
107. Clooney AG, Eckenberger J, Laserna-Mendieta E, et al. Ranking microbiome variance in inflammatory bowel disease: a large longitudinal intercontinental study. *Gut*. 2021;70(3):499–510. doi:10.1136/gutjnl-2020-321106
108. Huang J, Zhu X, Ma Y, et al. Machine learning in the differential diagnosis of ulcerative colitis and Crohn's disease: a systematic review. *Translat Gastroenterol Hepatol*. 2025;10:10. doi:10.21037/tgh-24-87
109. Rosenbacke R, Å M, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. *JMIR AI*. 2024;3:e53207. doi:10.2196/53207
110. Rynazal R, Fujisawa K, Shiroma H, et al. Leveraging explainable AI for gut microbiome-based colorectal cancer classification. *Genome Biol*. 2023;24(1):21. doi:10.1186/s13059-023-02858-4
111. Yu S, Li J, Ye Z, et al. Identification of a 10-species microbial signature of inflammatory bowel disease by machine learning and external validation. *Cell Regen*. 2025;14(1):32. doi:10.1186/s13619-025-00246-w
112. Al Radi ZMA, Prins FM, Colliv V, et al. Exploring the predictive value of gut microbiome signatures for therapy intensification in patients with inflammatory bowel disease: a 10-year follow-up study. *Inflamm Bowel Dis*. 2024;30(10):1642–1653. doi:10.1093/ibd/izae064
113. Massaro CA, Meade S, Lemarié FL, et al. Gut microbiome predictors of advanced therapy response in Crohn's disease: protocol for the OPTIMIST prospective, longitudinal, observational pilot study in Canada. *BMJ Open*. 2025;15(3):e094280. doi:10.1136/bmjopen-2024-094280

Journal of Inflammation Research

Publish your work in this journal

The Journal of Inflammation Research is an international, peer-reviewed open-access journal that welcomes laboratory and clinical findings on the molecular basis, cell biology and pharmacology of inflammation including original research, reviews, symposium reports, hypothesis formation and commentaries on: acute/chronic inflammation; mediators of inflammation; cellular processes; molecular mechanisms; pharmacology and novel anti-inflammatory drugs; clinical conditions involving inflammation. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>

Dovepress
Taylor & Francis Group