GUT MICROBES REPORTS
2026, VOL. 3, NO. 1, 2650247
https://doi.org/10.1080/29933935.2026.2650247

Taylor & Francis
by informa

REVIEW ARTICLE

OPEN ACCESS Check for updates

Precision nutrition through diet-gut microbiome interactions: Emerging insights driven by artificial intelligence, microbiome health metrics, and mechanistic modeling

Maria A. Barrera-Suarez[a], Conan Y. Zhao[b], Lioudmila V. Karnatovskaia[c] and Jaeyun Sung[d,e]

[a]Post-baccalaureate Research Education Program (PREP), Mayo Clinic Graduate School of Biomedical Sciences, Mayo Clinic, Rochester, USA; [b]Mayo Clinic Alix School of Medicine, Mayo Clinic, Rochester, USA; [c]Department of Internal Medicine, Division of Pulmonary, Critical Care Medicine, Sleep, and Allergic Diseases, Mayo Clinic, Rochester, USA; [d]Department of Quantitative Health Sciences, Division of Computational Biology, Mayo Clinic, Rochester, USA; [e]Microbiomics Program, Center for Individualized Medicine, Mayo Clinic, Rochester, USA

ABSTRACT

Diet-gut microbiome interactions drive substantial inter-individual variability in metabolic responses to food, a fact that challenges the efficacy of uniform dietary recommendations. To address this complexity, advances in multi-omics profiling, dietary assessment technologies, and host clinical phenotyping now generate high-resolution multimodal datasets. However, managing these vast amounts of data necessitates the integration of artificial intelligence (AI) and machine learning (ML) approaches. In this review, we first delineate the multimodal data landscape and its associated computational workflows. These range from the initial preprocessing of heterogeneous inputs (filtering, normalization, dimensionality reduction) to ML modeling strategies designed to address high dimensionality, sparsity, and compositionality through feature engineering and regularization. We then summarize core ML applications, including the classification of habitual dietary patterns from microbiome signatures, prediction of postprandial metabolic responses, responder stratification, and *in silico* simulation of dietary perturbations. Furthermore, recent randomized controlled trials demonstrate the tangible clinical potential of AI-guided personalization. Next, we highlight composite microbiome health metrics and diet-specific indices, such as GMWI2 and DI-GM. These tools are essential because they condense high-dimensional taxonomic profiles into interpretable wellness scores for monitoring diet-induced shifts. We subsequently examine genome-scale metabolic models and microbiome "digital twins" that mechanistically link dietary substrates to community metabolism and host-relevant metabolites. We also discuss emerging hybrid AI-mechanistic frameworks that enhance interpretability, biological plausibility, and scalability. Finally, we outline translational priorities—including the development of diverse longitudinal cohorts, standardized benchmarking, and clinically trustworthy AI—that are required to realize equitable, microbiome-informed precision nutrition.

ARTICLE HISTORY
Received 5 January 2026
Revised 25 February 2026
Accepted 18 March 2026

KEYWORDS
Precision nutrition; diet-gut microbiome interactions; artificial intelligence; machine learning; multi-omics profiling; composite microbiome health metrics; genome-scale metabolic models.

1. Introduction

The complex interplay among diet, the gut microbiome, and human physiology—three deeply interdependent systems—underlies the substantial inter-individual variation observed in metabolic and clinical responses to the same dietary interventions.[1,2] Growing evidence suggests that much of this variability is driven by differences in baseline gut microbial composition, establishing the microbiome as a key modulator of dietary efficacy.[3]

Although the adult gut microbiome exhibits relative stability,[4] it remains highly responsive to external influences, particularly diet.[5] As a result, dietary patterns exert selective pressures on gut microbial communities, reshaping their composition and metabolic potential.[6] For example, plant-based diets rich in dietary fiber foster expansion of saccharolytic taxa, thereby enhancing production of short-chain fatty acids (SCFAs) and other bioactive metabolites. In contrast, animal-based diets are often associated with

CONTACT Jaeyun Sung, PhD Sung.Jaeyun@mayo.edu Department of Quantitative Health Sciences, Division of Computational Biology, Mayo Clinic, 200 1st St. SW, Rochester, MN, 55905, USA

proteolytic microbial pathways and elevated levels of amino acid-derived metabolites.[5] These diet-induced changes occur across temporal and geographical scales, ranging from seasonal fluctuations observed in hunter-gatherer communities[7,8] to marked shifts following migration to industrialized settings.[9] Moreover, controlled studies in gnotobiotic mice[10,11] and human feeding trials[12,13] further demonstrate that acute dietary shifts can remodel gut microbial composition and function within as little as one day.

These rapid adaptations carry profound functional consequences, as gut microbes are active partners rather than passive residents. They contribute to host metabolic homeostasis and immune regulation,[14,15] maintain intestinal barrier integrity,[16] facilitate nutrient absorption,[17] and support the synthesis of essential vitamins and amino acids.[18,19] Consequently, disruption of microbiota composition or function (or "dysbiosis") has been implicated in the pathogenesis and progression of numerous metabolic, inflammatory, and even neuropsychiatric disorders.[10,15,20-23]

Within this framework, personalized nutrition extends beyond population-based dietary guidelines by tailoring recommendations to an individual's biological, metabolic, and lifestyle characteristics to improve health outcomes.[24] A major hurdle persists in resolving the complex, non-linear relationships between diet, microbiome, and host physiology. Traditional statistical techniques are frequently inadequate for handling the high dimensionality, heterogeneity, and multi-omic structure characteristic of modern datasets. Consequently, advanced computational approaches are increasingly used to identify hidden patterns, integrate diverse data types, and develop robust predictive models that extend beyond the limits of classical analyzes.[25] Among these, artificial intelligence (AI) offers a powerful framework for fusing high-dimensional data, from microbial and clinical markers to granular dietary information to inform personalized interventions and data-driven decision-making.[26-28]

This review synthesizes recent progress in applying AI and related computational tools to characterize diet-gut microbiome interactions and their implications for human health. We outline the current multi-omic and multimodal data landscape, examine key clinical and research applications, and address persistent challenges and limitations. Finally, we identify future priorities for embedding these approaches within precision nutrition frameworks. Collectively, the evidence highlights how AI and machine learning are reshaping our understanding of diet-gut microbiome dynamics and enabling evidence-based, personalized strategies to promote metabolic resilience and long-term health.

2. Multi-omics and multimodal data in diet-gut microbiome research

2.1. Multi-omics profiling of gut microbiota

In microbiome research, multi-omics datasets typically encompass four complementary layers: metagenomics, metatranscriptomics, metabolomics, and metaproteomics (Figure 1). Together, these approaches enable integrated profiling of microbial taxonomy, gene expression, metabolite production, and protein abundance. Although each omics layer targets a distinct molecular endpoint, the overall experimental workflow is conceptually similar. In all cases, biological samples are processed to selectively extract the molecules of interest—such as DNA, RNA, metabolites, or proteins—followed by measurement using high-throughput technologies. These procedures generate quantitative profiles describing the abundance and composition of genes, transcripts, metabolites, or proteins within each sample, which form the basis for downstream bioinformatic and statistical analyzes. Detailed reviews of multi-omics applications in microbiome research, including best practices for study design and analysis, are widely available.[29-31]

Genomic sequencing remains the cornerstone for characterizing the gut microbiome and is divided into two primary strategies: amplicon-based and shotgun metagenomic sequencing. Amplicon-based approaches target the bacterial/archaeal 16S rRNA gene or the eukaryotic 18S rRNA gene. These genes contain conserved regions (C1-C10) shared across microorganisms alternated with variable regions (V1-V9) that evolve over time. The conserved regions enable the design of broadly applicable primers, while the variable regions provide the taxonomic and phylogenetic signal used to distinguish microbial taxa. In contrast, Whole Genome Shotgun (WGS) metagenomics enables a more comprehensive assessment by untargeted sequencing of all DNA present in a sample. This approach achieves species- and often strain-level resolution,[32-34] enables direct reconstruction of microbial gene catalogs,[35,36] and facilitates *de novo* assembly of draft genomes.[37,38]

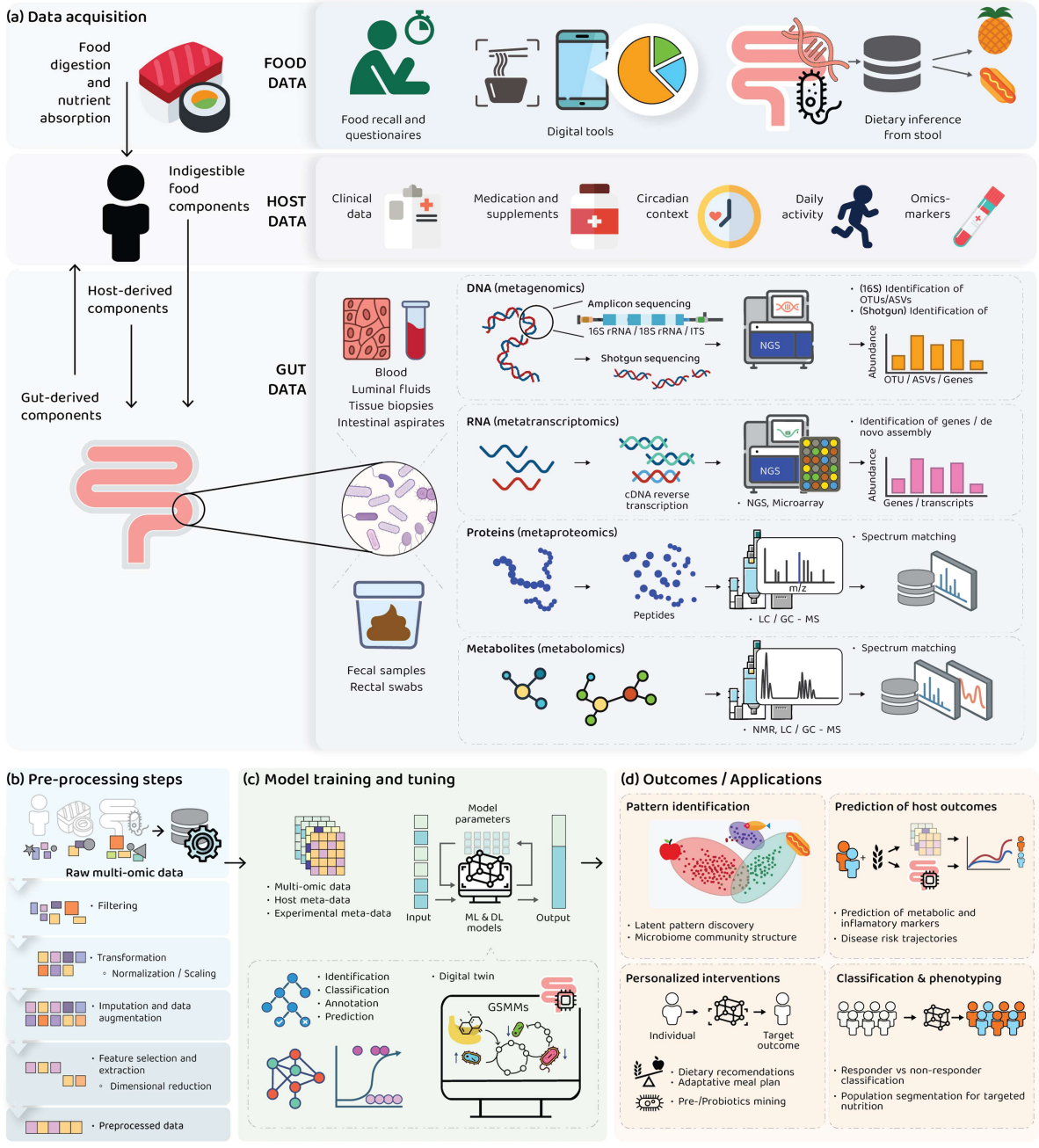

Figure 1. Multimodal data landscape and computational workflow in AI-driven diet-gut microbiome research. The schematic illustrates an end-to-end framework integrating dietary, host, and gut microbiome data to enable precision nutrition. (a) Data acquisition: The gut microbiome, diet, and host health are linked through a complex interplay. Dietary components are absorbed directly by the host or fermented by gut microbes, generating metabolites that influence host physiology. In turn, the host supplies endogenous substrates to the microbiome. (top) food data (self-reported recalls/questionnaires, digital tools, and emerging stool-based food-DNA inference); (middle) host data (clinical biomarkers, medication/supplement history, circadian context, daily activity, and omics-derived markers); (bottom) and gut microbiome multi-omics (metagenomics, metatranscriptomics, metaproteomics, metabolomics via next-generation sequencing (NGS) and LC/GC-mass spectrometry). (b) Preprocessing: Raw multi-omic and metadata inputs undergo filtering, transformation, normalization/scaling, imputation, data augmentation, feature selection/extraction, and dimensionality reduction. (c) Model training and tuning: multi-omic and multi-modal data are integrated into ML and DL models, enabling identification, classification, annotation, and prediction, as well as digital twin simulations based on genome-scale metabolic models (GSMMs). (d) Outcomes and applications: Trained models support latent pattern discovery and microbiome community characterization, prediction of host metabolic and inflammatory outcomes and disease risk trajectories, population stratification and phenotyping (e.g., responder vs non-responder classification), and personalized interventions, including adaptive dietary recommendations, meal planning, and targeted pre-/probiotic strategies.

Both approaches have distinct advantages and limitations.[30,39] Amplicon sequencing is comparatively low-cost, scalable to large cohorts, and supported by well-stablished pipelines, making it suitable for profiling microbial community composition across many samples.[40] However, it is susceptible to primer bias and PCR amplification artifacts, exhibits uneven detection of taxa with variable gene copy numbers, and provides limited insight into functional potential.[40-44] In contrast, WGS enables functional pathway inference, detection of viruses and fungi, and improved taxonomic resolution. Because all genetic material in a sample is sequenced, host DNA contamination represents a limitation not only at the analytical level but also at the sample level. High host-to-microbial DNA ratios—particularly in low microbial biomass samples—can substantially reduce microbial signal and, in some cases, render specimens unsuitable for sequencing. In addition, WGS metagenomics incurs higher sequencing costs, greater computational demands, dependence on reference databases for taxonomic and functional annotation, and analytical challenges related to uneven coverage, and complex assembly and binning steps.[45-47] Importantly, both amplicon-based and WGS data are inherently compositional, such that observed relative abundances are constrained by total sequencing depth. To address this limitation, a range of compositional data analysis techniques, normalization strategies, and differential abundance frameworks have been developed specifically for this type of data.[48]

While genomic sequencing captures the repertoire of genes present in a microbial community, it does not reflect transcriptional activity. Metatranscriptomics overcomes this limitation by profiling RNA to characterize time- and condition-dependent gene expression. Despite its promise, metatranscriptomics faces multiple obstacles, such as RNA instability, challenges associated with ribosomal RNA depletion, ongoing host RNA contamination, variable extraction efficiency, and difficulties in assigning transcripts to specific taxa due to gene homology.[49-51] As a result, careful experimental and analytical protocols are essential for reliable data replication and interpretation.

Metaproteomics adds an additional layer of resolution by quantifying proteins, thereby linking transcript abundance to functional output shaped by post-transcriptional regulation, translational efficiency, and protein turnover. Metaproteomic methods based on mass spectrometry (MS) further facilitate detection of microbial enzymes, structural proteins, and host-microbial interaction factors, yielding insights into active metabolic pathways and functional states within the gut microbiome.[52-56] However, the application of metaproteomics in gut microbiome research remains limited by the proteome's broad dynamic range, incomplete reference databases, peptide-to-taxon ambiguity, and contamination from host proteins.[57]

Metabolomics captures the diverse repertoire of small-molecule metabolites produced by the gut microbiota using MS-based platforms or nuclear magnetic resonance spectroscopy. These metabolites exert wide ranging effects on host physiology, including maintenance of intestinal barrier integrity, immune regulation, and control of lipid and glucose metabolism.[58-61] A key strength of metabolomics is its ability to reveal functional outputs that are not strictly linked to microbial taxonomic composition, reflecting functional redundancy and metabolic plasticity within microbial communities.[62,63] Accordingly, stable microbial composition does not necessarily imply stable metabolic activity, and clinically relevant differences can occur without clear changes in community structure.[64,65] Despite these insights, metabolomic analyzes remain challenging to interpret, as metabolite profiles reflect integrated signals from microbial activity, host metabolism, diet, and environmental exposures,[66] and are further affected by technical variability and incomplete annotation in current platforms and databases.[67]

Although each omics layer yields valuable insights, none provides a comprehensive perspective in isolation. Integrative analyzes are limited by differences in scale, data sparsity, temporal resolution, and noise. Moreover, incomplete and inconsistent metadata in public repositories can compromise reproducibility and introduces the risk of confounded cross-omics associations, underscoring the need for FAIR (Findable, Accessible, Interoperable, Reusable)-compliant metadata standards.[68]

2.2. Dietary intake assessment methods

Accurate dietary assessment is a fundamental yet enduring challenge in diet-gut microbiome research. Conventional approaches rely primarily on self-reported instruments, including food frequency questionnaires (FFQs), 24-hour dietary recalls, and food diaries, which are scalable but susceptible to recall errors,

reporting biases, portion misestimation, and high participant burden.[69,70] Their limitations are further affected by the vast diversity of global dietary patterns. Although major nutrient databases such as the U.S. Department of Agriculture's FoodData Central[71] and Canadian Nutrient File[72] now catalog hundreds of thousands of food items, many bioactive food components remain poorly characterized or unquantified.[73] More objective methods, such as weighed records or image-assisted logging, can improve measurement precision but are labor-intensive and impractical for large cohorts.

AI is increasingly transforming dietary assessment. AI-based food image recognition systems—including NutriNet,[74] goFOOD,[75] the Mobile Food Record,[76] and related smartphone apps[77,78]—automate meal segmentation, food identification, portion estimation, and nutrient mapping. These tools reduce participant burden and allow more frequent assessment of dietary intake. Additionally, food diary data can be mapped to nutrient composition tables to generate structured nutrient profiles,[79] although cross-database inconsistencies, regional food variability, and labeling discrepancies continue to limit comparability.

Biomarker-based validation provides an independent means of assessing dietary intake. Established markers include urinary nitrogen for estimating protein intake,[80] skin carotenoid levels for fruit/vegetable consumption,[81] and stable carbon and nitrogen isotope ratios measured in hair or blood for animal protein[82] or added sweeteners.[83] Despite its promise, metabolomics-based dietary assessment is restricted by the lack of specificity of many biomarkers.[84] For instance, metabolites such as carotenoids or vitamin C are shared across numerous fruits and vegetables, making it difficult to unambiguously attribute signals to individual food or dietary components.[85]

An innovative approach to overcome the challenges inherent to self-reported dietary assessments involves using fecal metabolomics,[86] metaproteomics[87] or sequencing of food-derived DNA from stool to infer dietary intake directly. Computational tools like FoodSeq[88] (targeted metabarcoding of food-specific markers) and MEDI[89] (untargeted detection of food-derived reads in shotgun data) identify residual plant and animal DNA originating from consumed foods, providing an objective, sample-linked readout of dietary exposure. Complementing these molecular approaches, AI-based methods such as METRIC (Microbiome-based Nutrient Profile Corrector) use gut microbiome composition to statistically reduce random error in self-reported nutrient intake profiles.[27] Rather than replacing dietary records, METRIC denoises them, leveraging microbiome signals as a biological reference to improve the reliability of estimated nutrient intake. These approaches help detect foods or nutrient patterns that may be missed or underreported, providing dietary information that is often unavailable in retrospective analyzes.[90] Yet several important limitations remain, including variable persistence of food-derived DNA through digestion, incomplete reference databases, and inability to quantify portion size or habitual patterns. These emerging techniques are most effective when integrated with traditional self-reported data and established biomarkers.

2.3. Host physiological and clinical data types

Beyond microbiome composition and dietary intake, host physiological and clinical data provide essential context for understanding individualized responses to diet. These variables encompass a wide spectrum, from high-dimensional host omics to standard clinical and phenotypic measures—such as anthropometrics (e.g., BMI, waist-to-hip ratio), blood biomarkers, medication history, and lifestyle factors (e.g., physical activity, sleep). Integrating host features with microbiome and dietary information creates a more comprehensive view of the inter-individual variability that ultimately shapes dietary outcomes.[91]

Models that incorporate host variables alongside diet and microbiome profiles consistently outperform those based on diet or microbiome data alone, as demonstrated, for example, in the prediction of postprandial glycemic responses (PPGR), where inclusion of host variables improves the accuracy of predictive algorithms.[91] Host factors directly reflect underlying metabolic status; individuals with insulin resistance or elevated fasting glucose, for instance, typically exhibit larger post-meal glucose excursions.[92] Demographic attributes (age, sex) and lifestyle covariates (physical activity, medication use) further modulate these responses and are therefore routinely included as key predictors in algorithmic frameworks.

Despite their clear utility, host-derived measurements share many of the analytical challenges seen in microbiome and dietary data, including heterogeneity in scale and type, measurement error, mismatched

temporal resolution, and substantial inter-individual variability. Addressing these substantial challenges is essential to support reliable multimodal integration and realizing truly personalized nutritional insights.

3. Machine learning and deep learning applications in diet-gut microbiome interactions

Diet-gut microbiome research increasingly relies on computational frameworks that draw from machine learning, deep learning, and mechanistic modeling. Given the interdisciplinary nature of this field, key terminology from artificial intelligence and predictive modeling is briefly summarized in Box 1 to provide conceptual clarity for readers from diverse backgrounds.

Diet-gut microbiome research yields high-dimensional and multimodal datasets characterized by non-linear and context-dependent interactions among diet, microbial communities, and host physiology. Although traditional statistical techniques are well suited for hypothesis testing, they often struggle to model complex biological systems, particularly when developing generalizable predictors across diverse individuals. Many classical models rely on linear assumptions, treat features independently, and require pre-specified hypotheses, which limits their ability to capture nonlinear relationships and complex interactions.[93] Moreover, these approaches are sensitive to high dimensionality, sparsity, and collinearity, and they often fail to account for substantial inter-individual variability in the microbiome.

As a result, the domain has pivoted toward AI solutions, specifically machine learning (ML) approaches. ML refers to a class of algorithms that learn patterns directly from data to make predictions or decisions without being explicitly programmed with fixed rules. These algorithms ranging from ensemble-based methods to deep learning (DL) architectures excel at fusing heterogeneous data modalities and maximizing performance on held-out samples.[29] Within this framework, DL represents a specialized subfield of ML that uses multi-layer neural networks to learn hierarchical feature representations directly from high-dimensional inputs.[94,95] The key difference lies in DL's ability to discover intricate structures in complex datasets by using backpropagation to adjust internal parameters, thereby computing representations in each layer from the previous one.[94,96] This process enables end-to-end learning of hierarchical representation directly from raw data, without the need for manual feature engineering. Broader applications of ML in microbiome science have been extensively reviewed elsewhere.[25,97,98]

In diet-gut microbiome research, ML has enabled several transformative applications (summarized in Table 1). First, supervised and unsupervised ML models have successfully classified habitual dietary patterns directly from taxonomic or functional gut microbiome profiles,[99-102] revealing reproducible microbial signatures of long-term intake. For example, gut microbiomes differentiated vegan, vegetarian, and omnivore diets (mean AUC = 0.85), showing consistent taxonomic shifts: omnivores were enriched in bile-tolerant and protein-fermenting taxa, vegans in fiber-degrading and SCFA-producing bacteria, and vegetarians in dairy-associated microbes.[99] Second, integrative models combining dietary logs with baseline microbiome data reliably predict individual metabolic biomarkers. These include postprandial metabolic responses—most robustly glycemic excursions[2,91,103-108] but also lipidomic[2,107] and inflammatory markers in select contexts. These models consistently outperform diet-only or microbiome-only baselines by capturing person-specific interactions that drive variable responses to identical meals.[91,103]

Translationally, microbiome-informed ML has generated personalized dietary recommendations, optimizing macronutrient ratios or food choices for metabolic outcomes.[91,104,105,116] Clinical trials based on this approach have demonstrated benefits in glycemic control and related endpoints, although superiority over standardized guidance is context-dependent and not universal across all populations or outcomes.[106] Additionally, ML stratification of baseline gut microbiomes can prospectively identify likely responders vs. non-responders to dietary interventions,[122-125,127] supporting precision enrollment and resource allocation.

More recently, representation-learning and multi-task DL frameworks have begun to model underlying microbial dynamics and their downstream functional consequences. A prominent example is the Metabolite-Response Predictor using Coupled Multilayer Perceptrons (McMLP),[130] which jointly predicts post-intervention microbiome reconfiguration and resulting metabolomic shifts. By chaining two neural networks—one forecasting community composition from baseline microbiome, metabolome, and diet, and the second mapping the predicted community to metabolite abundances—McMLP explicitly encodes microbe-microbe and microbe-metabolite dependencies. Across multiple dietary intervention cohorts, it outperformed uncoupled baselines, particularly for health-relevant microbial metabolites. Such architectures

Table 1. Representative applications of machine learning in diet-gut microbiome research.

Research theme	Representative application	Machine learning approach	Data type	Key findings	References
Habitual diet shapes gut microbiome composition	Identification of diet-associated microbial taxa and functional pathways across large population cohorts	• RF classification and regression [87,99,107] • Regularized regression [109] • Unsupervised clustering [110,111110,111]	Host data/dietary data/ Metagenomics [99,107,109,110] (or Metabolomics [111]) (or Metaproteomics [87])	Long-term dietary patterns strongly determine gut microbiome composition. Diets high in processed and animal-derived foods are associated with enrichment of bile-tolerant and pro-inflammatory taxa [99,107,109,110], alongside increased endotoxin synthesis pathways [109,110]. In contrast, plant-based and vegan patterns correlate with higher abundances of fiber-degrading, short-chain fatty acid-producing taxa and butyrate producers [99,107,109,110] which are linked to favorable cardiometabolic profiles.	[87,99,107,109-111]
	Classification and prediction of dietary patterns based on gut microbiome composition	• RF classification and regression [86,99-102,112] • Unsupervised clustering [113]	Host data/dietary data/ Metatranscriptomics [100] (or Metagenomics [99,101,102] or Metabolomics [86,112,113])	Habitual dietary patterns can be robustly classified from gut microbiome composition, (e.g., distinguishing vegan, vegetarian, and omnivorous diets [99,100], diet quality.[113]) Diet quality classification also achieves high accuracy, with high-fiber diets enriched in butyrate-producing taxa (e.g., Lachnospiraceae) and Western/low-quality diets characterized by bile-tolerant and protein-fermenting bacteria (*e.g., Bilophila, Blautia*) [101]. In animal models, machine learning reliably discriminates high-fat diets and protein sources, identifying *Lactococcus spp.* as a key indicator and separating plant- from animal-protein diets, with plant protein linked to greater microbial diversity and fiber-degrading taxa, and animal protein to protein-fermenting Clostridia [102].	[86,99-102,112-114]
Prediction of metabolic biomarkers from diet and microbiome data	Prediction of postprandial glycemic responses (PPGRs)	• GBDT [91,103-106,115] • RF classification and regression [2,107] • Deep neural network [108]	Host data/dietary data/ Metagenomics [91,103-108] (or Metatranscriptomics [115])	Across diverse populations and metabolic states, postprandial glycemic responses can be accurately predicted by integrating dietary intake with gut microbiome features. Models incorporating diet and microbiome outperform diet-only or carbohydrate-based approaches, capturing substantial inter-individual variability in responses to identical foods [2,91,103]. Microbiome features linked to plant-rich, fiber-dense diets are associated with more favorable PPGR profiles, whereas	[91,103-108,115]

(*Continued*)

Table 1. (*Continued*)

Research theme	Representative application	Machine learning approach	Data type	Key findings	References
				bile-tolerant or dysbiotic signatures predict higher glycemic excursions [107]. Incorporating microbiome data further improves PPGR prediction accuracy in pre-diabetic [104], type 2 diabetes mellitus [105,108] and pregnant women with or without gestational diabetes [106].	
	Prediction of lipid, inflammatory, cardiometabolic risk biomarkers or another host outcome	• RF regression and classification [2,107]	Host data/dietary data/ Metagenomics [2,107]	Machine-learning models indicate that meal composition plays a relatively smaller role in metabolic responses, with fat-induced triglyceride excursions largely driven by inter-individual biological variation. These models achieve moderate accuracy in predicting postprandial triglyceride rises, [2] as well as total lipid, cholesterol, and triglyceride levels across multiple lipoprotein fractions in both fasting and postprandial states [107].	[2,107]
Personalized dietary interventions	Microbiome-guided dietary recommendations to minimize or maximize output	• GBDT [91,104,105,116-118] • Proprietary ML model [119]	Host data/dietary data/ Metatranscriptomics [119] (or Metagenomic [91,104,105,116-118])	Machine learning-guided dietary personalization has been shown to reduce postprandial glycemic responses relative to control diets in multiple populations [91,104,105], as well as triglycerides [118], inflammatory markers [91], hepatic fat content [117], symptom score reduction [119], and HbA1c [118,120] levels. However, not all trials demonstrated superiority of personalized diet over standard dietary interventions [116,121].	[91,104,105,116-119]
	Identification of microbiome-defined responder and non-responder to dietary interventions	• RF regression and classification [122-126] • GBDT [127] • Regularized regression [128] • Unsupervised methods [123]	Host data/dietary data/Metagenomics [122-127] (or Metabolomics) [128]	Baseline gut microbiome composition was used to stratify and cluster individuals prior to dietary intervention, enabling prediction of responders vs. non-responders for weight loss [122,124,125], hepatic fat content [129], and glycemic outcomes [122], including reductions in HbA1c [127], fasting plasma glucose [127], PPGRs and intestinal function [123,126]. These response-defined clusters were characterized by distinct dominant microbial taxa and microbiome structure.	[122,123,125-128]

Table 1. (*Continued*)

Research theme	Representative application	Machine learning approach	Data type	Key findings	References
In silico modeling of microbiome-host response	Simulation of host metabolic and microbiome shifts to dietary perturbations	• Deep multilayered neural networks [130,131]	Host data/dietary data/ Metagenomics [130,131] (or Metabolomics [130])	*In silico* models can predict host metabolic responses and microbiome shifts to dietary perturbations. ML-based models were able to learn latent representations of gut microbiome community structure and capture predictable shifts induced by dietary interventions [131]. By explicitly modeling diet-associated changes in microbial composition, these learned representations were then used to predict host metabolomic responses [130].	[130,131]

Abbreviations: RF, random forest; GBDT, gradient-boosted decision trees; PPGR, postprandial glycemic response; ML, machine learning.

illustrate the capacity for DL to bridge structural and functional layers, moving beyond static prediction toward simulation of dietary perturbations.

While ML has shown great promise for predictive applications in diet-gut microbiome research, its practical use is often constrained by familiar challenges, notably interpretability, the need for large training datasets, and the requirement for rigorous model evaluation.[25,132] As most ML models capture statistical associations between inputs and responses without explicitly representing underlying biological mechanisms (associative signals rather than causal), their decision-making processes are often described as "black-box" predictions.[133] This lack of transparency becomes particularly problematic in clinical or translational settings, where mechanistic interpretability is essential for building trust, facilitating regulatory approval, and enabling actionable biological insights. Furthermore, ML models tend to require dense datasets, are vulnerable to overfitting, and prone to dataset sensitivity, making rigorous parameter regularization, careful (nested) cross-validation, and systematic, external benchmarking essential.

To enhance transparency, the community has increasingly adopted post-hoc interpretability methods. These include feature attribution techniques (e.g., permutation importance,[134] LIME[135]) and SHAP values,[136,137] which are broadly applicable across ML models to quantify the contribution of individual features (e.g., specific taxa, pathways, or dietary components) to predictions; saliency mapping[138] for gradient-based sensitivity analysis; and attention weights,[139] which are particularly prominent in DL architectures to highlight relevant input elements within neural network processing. Comprehensive strategies for interpretable ML in biological contexts are discussed in depth elsewhere.[140]

4. Advances in gut microbiome-informed precision nutrition

Precision nutrition marks a fundamental shift from conventional "one-size-fits-all" dietary guidelines toward tailored recommendations that leverage an individual's unique biological profile, gut microbiome, and lifestyle factors.[24,141,142] By combining diverse datasets (e.g., multi-omics, dietary logs, and clinical phenotypes) via advanced computational pipelines, this approach enables mechanism-driven interventions grounded in host-microbe dynamics and physiological variability.[27] The overarching goal of precision nutrition is to fine-tune metabolic outputs, mitigate disease risk, and promote sustained health through data-centric personalization.

A prominent application of machine learning in precision nutrition is the prediction of individual clinical responses to dietary intake. Likewise, these trained models can be used to generate AI-guided dietary recommendations designed to achieve specific clinical objectives. Previous studies have demonstrated that identical dietary interventions often produce highly variable responses across individuals, a phenomenon partly explained by differences in gut microbiome composition.[143] In response to this variability, early efforts sought to integrate microbiome data into predictive frameworks, most notably in the pioneering study by Zeevi et al.[91] In this work, dietary, clinical, behavioral, and gut microbiome features were integrated into a gradient-boosted ML model to predict individual post prandial glucose responses (PPGRs). By jointly modeling host and microbial factors, the framework far surpasses traditional carbohydrate-centric and population-level prediction approaches. A follow-up randomized controlled trial (RCT) validated the model's utility, showing that algorithm-derived "good" diets minimized glucose spikes.[105] Importantly, foods predicted to be optimal or suboptimal were highly individual-specific, such that the same food could elicit beneficial effects in one individual but adverse responses in another. Other studies have replicated and validated these findings.[103,105,115,144]

Building on these foundations, large-scale cohorts have refined and expanded predictive scope. The PREDICT 1 study extended modeling beyond glycemic to lipidomic and insulinemic responses using a random forest–based approach.[2] In this study, gut microbiome composition explained a substantial portion of individual variability in PPGRs, exceeding the contribution of macronutrient content alone. Moreover, specific microbial profiles associated with plant-rich diets were linked to more favorable cardiometabolic outcomes, including improved fasting glucose, postprandial glycemic and lipid responses, and reduced inflammation.[107] Together, these findings highlight the central role of the gut microbiome in shaping individual dietary responses. More recently, DL architectures have been developed to integrate dietary intake, continuous glucose monitoring (CGM) data, clinical variables, and gut microbiome profiles to predict PPGRs in individuals with type 2 diabetes.[108] These models demonstrate robust predictive

performance, particularly improving predictions in low-carbohydrate responders, and achieve accuracy comparable to conventional machine-learning approaches.

The field has progressed from *in silico* proofs-of-concept to rigorous RCTs assessing AI-guided diets' clinical efficacy. In adults with prediabetes, a RCT compared a standard Mediterranean diet with a personalized postprandial-targeting (PPT) diet.[144] The personalized arm showed significantly greater reductions in CGM-derived hyperglycemia and HbA1c, with benefits sustained through 12-month follow-up. In a companion study, individuals with newly diagnosed type 2 diabetes who underwent the PPT diet intervention showed improved PPGRs control, reductions in HbA1c, fasting glucose, and triglycerides, and high rates of diabetes remission.[105]

Similar findings have been reported in gastrointestinal disorders. In an RCT involving patients with irritable bowel syndrome (IBS), an AI-assisted, microbiome-based personalized diet was compared with a standard low-FODMAP diet.[121,145] Although both interventions produced meaningful improvements in symptoms and quality of life, the microbiome-targeted diet was associated with distinct shifts in gut microbial composition and diversity. In non-alcoholic fatty liver disease (NAFLD), AI-guided, microbiome-informed dietary interventions have likewise demonstrated clinical promise. In a randomized study, personalized diets designed using metagenomic profiling and ML, combined with targeted probiotic supplementation, resulted in greater reductions in liver fat and improvements in microbial and inflammatory markers compared with standard dietary advice.[117]

Despite promising outcomes, the efficacy of AI-guided personalized diets is not universal and appears context-dependent. For instance, in an RCT involving adults with obesity and abnormal glucose metabolism,[116] participants were assigned to either a standardized low-fat diet based on the Zeevi et al.[91] algorithm designed to predict postprandial glycemic responses. Both groups received equivalent behavioral counseling in the trial, in which evaluated weight loss as the primary outcome. At 6 months since initiation of the trial, the personalized approach did not yield additional weight-loss benefits compared to the standardized diet. These findings suggest that algorithms optimized for specific endpoints, such as glycemic control, may not translate effectively to other outcomes like weight loss, further suggesting the need for outcome-specific predictive models.

Extending beyond glycemia, dietary impacts on health often manifest through microbially generated metabolites. Evaluating these effects is a key aspect of nutritional research, as changes in the concentrations of specific health- or disease-associated metabolites provide a direct measure of an intervention's impact. However, direct metabolomics, however, is resource-intensive, technically demanding, and susceptible to contextual artifacts. To circumvent this, researchers have developed computational frameworks that predict diet-induced metabolite dynamics from microbiome composition, thereby establishing a functional bridge between microbial communities and host metabolic outcomes. Beyond the McMLP discussed above, a variety of ML approaches have been proposed for inferring metabolite profiles from paired microbiome-metabolome datasets. These data-driven methods operate independently of predefined biochemical pathways or reference databases and encompass techniques such as Elastic Net regression,[146] sparsified neural encoder-decoder networks,[147] multilayer perceptrons,[148] word2vec embeddings,[149] neural ordinary differential equations,[150] and mixture-of-experts models with variational Gaussian processes.[151]

Despite these strides, AI applications in diet-gut microbiome-host modeling face persistent hurdles. Models often degrade under domain shifts (e.g., cohort, cuisine, or sequencing variations) and rarely incorporate temporal elements like circadian effects or meal sequencing.[152] Although this may change as precision nutrition research increasingly emphasizes not only *what* people eat but also *when* they eat. Consistent with this direction, the NIH 2020–2030 Strategic Plan for Nutrition Research[153] identifies circadian rhythm as an addressable component of precision nutrition, specifically highlighting chrononutrition and meal-timing as emerging priorities. Most models also lack native uncertainty estimation or mechanistic transparency, undermining clinical trust and deployment. Emerging hybrids address this by blending deep representation learning with probabilistic inference (e.g., mixture-of-experts plus variational Gaussian processes[151,154]) and attribution methods to dissect microbe-metabolite contributions[148,155]—fostering interpretable insights into food-microbe-metabolite cascades. Looking ahead, practical translation hinges on closed-loop optimizers that iteratively design meals/timings for multi-objective metabolic gains, with safeguards for safety, practicality, and equity.

5. Mechanistic modeling of gut microbiome community metabolism

Complementing data-driven statistical and ML methods, genome-scale metabolic models (GSMMs) offer a mechanistic framework for mapping dietary inputs to microbial community metabolism. Well-established in systems biology,[156] GSMMs have become increasingly central in microbiome studies for simulating metabolic interactions in diverse gut consortia.[157] These models represent microbial species or strains as stoichiometrically balanced networks, incorporating thousands of reactions, transporters, and biomass objectives. Aggregated into multi-species community models,[93,158-160] they capture ecological dynamics such as cross-feeding, competition for resources, and collective metabolite outputs—SCFAs, amino acid derivatives, and bile acid conjugates. In-depth reviews of GSMM tools and resources tailored to gut microbiome applications are available elsewhere.[161]

Personalization begins with aligning an individual's metagenomic or amplicon profile to strain-level reconstructions from databases like AGORA,[162] AGORA2,[163] or the Virtual Metabolic Human (VMH) collection,[164] and by constraining the models with subject-specific dietary intake data. Nutrient databases and luminal availability estimates convert food intakes into uptake bounds, thereby defining the substrate pools and flux capacities for the modeled community. Constraint-based optimization techniques—most commonly flux balance analysis (FBA), flux variability analysis (FVA), or dynamic FBA (dFBA)—then compute feasible metabolic flux distributions, including exchange fluxes that predict metabolite production or consumption.[165,166]

GSMMs have significantly advanced mechanistic insights into dietary influences on gut microbial metabolism. Several dedicated computational frameworks—including COMETS,[167] BacArena,[168] dOptCom,[169] MatNet,[170] DyMMM,[171] MCM,[172] and the CASINO[173] toolbox—have been developed to investigate diet-gut microbiome interactions. These tools typically integrate multiple species- or strain-level GSMMs into microbial community-scale metabolic models (MCMMs), enabling simulation of community-wide metabolic adjustments under varying nutritional conditions.

Early translational efforts yielded encouraging results. CASINO, for example, generated personalized microbiota models for 45 overweight or obese individuals, and linked dietary composition to microbial metabolic capabilities along the diet-microbiota-host axis.[173] The models predicted marked shifts in amino acid fermentation and SCFA production in response to dietary changes, with simulations validated against paired fecal and plasma metabolomics.

More recent studies have shown that MCMMs derived from individual gut metagenomes can accurately predict personalized SCFA production across dietary, prebiotic, and probiotic interventions. In one large-scale example,[174] microbial community-scale models were constructed for hundreds of participants by integrating AGORA[162] strain-level reconstructions with MICOM,[175] a scalable platform for simulating mixed communities under defined nutritional constraints. Each model incorporated subject-specific metagenomic abundances and diet-derived substrate availability, yielding individualized SCFA flux predictions across multiple dietary intervention scenarios. Predicted acetate, propionate, and butyrate production correlated strongly with measured fecal SCFA levels and host cardiometabolic/immunological biomarkers. Notably, these personalized MCMMs were further leveraged to design targeted interventions, identifying for each individual the fiber types or probiotic strains predicted to maximally boost SCFA output.

Collectively, these developments herald an emerging paradigm in which MCMMs serve as individualized "digital twins" of the gut ecosystem. Such digital twins can be perturbed *in silico* to evaluate alternative diets, prebiotics, probiotics, or combination therapies, and calibrated against metatranscriptomic or metabolomic data to refine parameter estimates. GSMMs integrate naturally with AI workflows, as derived features such as pathway fluxes and exchange rates provide biologically grounded priors for supervised ML, improving interpretability and reducing artifacts arising from compositional.[176] In contrast, neural-mechanistic hybrids models[177] replace FBA with differentiable approximations, enabling phenotype inference for cohort-scale analyzes. In all, GSMM-based digital twins may emerge as a hypothesis-generating counterpart to data-driven ML.

Despite these strengths, key limitations remain, including steady-state assumptions, uncertain objective functions, incomplete metabolic reconstructions, and limited incorporation of regulatory dynamics or host-microbe co-metabolism.[178,179] Objective functions, often defaulted to biomass maximization, may not

reflect real-world microbial priorities like survival under nutrient limitation or cooperative behaviors in communities, introducing biases in flux estimates. Metabolic reconstructions remain incomplete, with gaps in pathway annotations—particularly for uncultured or understudied gut taxa—resulting in orphan reactions or erroneous inclusions from databases like KEGG or MetaCyc. Regulatory dynamics, including gene expression controls and post-transcriptional modifications, are poorly incorporated, decoupling models from omics realities where transcript-protein-flux relationships are non-linear. Host-microbe co-metabolism is similarly underrepresented, neglecting bidirectional exchanges like mucosal signaling or immune modulation that influence dietary outcomes. Furthermore, computational demands escalate for community-scale models, where solving large linear programs becomes intractable for high-resolution personalization or iterative sensitivity analyzes.

Looking ahead, the true power lies in hybrid AI-GSMM paradigms, where machine learning addresses core limitations to propel precision nutrition forward. Neural surrogates and tree-based ensembles can approximate FBA solvers, accelerating computations from hours to seconds for cohort-scale simulations and enabling real-time uncertainty quantification through ensemble variance or Bayesian priors.[177,180] Multi-task models fuse GSMM-derived fluxes as biologically informed features into predictive pipelines,[181] enhancing interpretability (e.g., via SHAP attribution on pathway contributions[182]) and mitigating compositional biases in microbiome data. Conversely, GSMM constraints can regularize ML objectives, imposing stoichiometric feasibility to curb overfitting and yield causally plausible predictions.[176] In diet-gut microbiome research, such synergies promise scalable, personalized tools—optimizing prebiotic/fiber blends for SCFA maximization or simulating host responses to novel diets[183]—ultimately bridging empirical patterns to mechanistic precision for metabolic health.

6. Beyond *α*-diversity: Composite and diet-specific metrics of gut microbiome health

Conventional α-diversity metrics, which quantify microbial community richness and evenness, have long served as proxies for dysbiosis but show inconsistent associations with host health outcomes. These inconsistencies arise because α-diversity reflects overall community structure without capturing the direction, magnitude, or functional implications of finer-scale compositional changes.[184] Moreover, dysbiosis lacks a precise definition owing to the absence of a universal "healthy" reference microbiome and extensive inter-individual variation; a healthy gut microbiome is not represented by a single configuration but rather by multiple configurations linked to host well-being.[185] Conceptually, dysbiosis represents a deviation from configurations typically associated with well-being. Effective quantification therefore demands reference-based indices that compare individual profiles against empirically validated healthy or disease-linked benchmarks, rather than relying solely on diversity summaries.

To overcome these limitations, ML approaches have allowed the formulation of composite microbiome health metrics that condense thousands of microbial features into a single, interpretable score reflecting host microbiome health. In dietary research and precision nutrition, these reference-based metrics make microbiome changes more interpretable and actionable. However, they are indicative rather than diagnostic and should not be interpreted as clinical endpoints. Their utility in dietary optimization is therefore indirect, serving primarily to benchmark microbiome states and track diet-associated shifts rather than to define health outcomes per se. Leading examples include the Gut Microbiome Health Index (GMHI)[186] and its advanced successor, the Gut Microbiome Wellness Index 2 (GMWI2),[187] as well as the Health Index with Principal Component Analysis (hiPCA).[188] Representative applications of these metrics are summarized in Table 2.

The GMHI,[186] trained on 4,347 shotgun metagenomes from 34 multi-cohort studies, computes a log-ratio of health-enriched vs. health-depleted taxa to generate a disease-agnostic risk estimate, far outperforming α-diversity metrics in classification tasks. This framework was subsequently refined into GMWI2, which expanded the training dataset to 8,069 metagenomes across 54 studies, replacing the log-ratio with Lasso-penalized logistic regression across multiple taxonomic ranks.[187] GMWI2, anchored in the presence/absence of disease-associated microbes, provides a continuous wellness score reflecting the degree of association with health.

Beyond binary case-control classification, GMWI2 has demonstrated sensitivity in longitudinal and interventional contexts. In dietary trials, it detected a rapid decline in gut health within two days of a fiber-

Table 2. Representative applications of composite microbiome health metrics.

Index	Application	Key findings
Gut Microbiome Health Index (GMHI) [186]	Characterize gut microbiome in breast cancer patients before and after chemotherapy [189]	GMHI detected significant deterioration in gut microbiome health following chemotherapy in breast cancer patients, with lower post-treatment GMHI scores indicating increased dysbiosis.
	Evaluate gut microbiome health across different microbiome states (defined as distinct community configurations) [190]	GMHI varied systematically across distinct gut microbiome community configurations, with *Clostridiales*-enriched states exhibiting higher GMHI values and *Bacteroides*-dominated low-diversity states showing lower GMHI.
	Quantify probiotic treatment in ulcerative colitis [191]	Probiotic supplementation significantly increased GMHI and reduced the microbial dysbiosis index in ulcerative colitis patients. Improvements in GMHI paralleled reductions in disease activity and inflammatory burden.
	Quantify whether puerarin treatment restored the gut microbiome from a high-fat-diet-induced dysbiotic state toward a healthier overall microbial profile [192]	Puerarin treatment restored GMHI in high-fat-diet-fed mice, reversing diet-induced dysbiosis despite minimal changes in alpha diversity.
	Assess microbiome health in relation to environmental factors (including diet) in a large Dutch population [193]	GMHI scores were significantly higher in healthy than in diseased participants and correlated with shared disease signatures. Diet emerged as a key factor, with plant-based patterns associated with health-enriched taxa (e.g., SCFA producers) and Western diets linked to inflammatory profiles and lower scores.
	Quantify how dietary adenine affects the gut microbiota [194]	Dietary adenine induced significant, dose-dependent reductions in GMHI across generations, indicating worsening gut microbiome health associated with elevated purine intake.
	Quantify the prebiotic effects of oligosaccharides on gut microbiome wellness using in vitro fecal fermentation [195]	Prebiotic treatment (FOS, GOS, XOS, IN, 2FL) resulted in positive GMHI values, indicating enhanced relative abundances of health-prevalent species (*e.g., Bifidobacterium adolescentis, Bifidobacterium catenulatum*) and reduced health-scarce species. Controls without substrate showed negative GMWI, reflecting a dysbiotic state. GMWI outperformed the traditional prebiotic index (PI) and α-diversity metrics in evaluating dietary intervention impacts on microbiome composition and health.
	Characterize the impact of kefir administration on the gut microbiome in critically ill adults [196]	GMHI improved significantly after kefir administration in ICU patients, indicating enhanced gut microbiome health despite no increase in α-diversity and ongoing antibiotic exposure. Kefir was safe and feasible, with variable engraftment of kefir-derived species.
Gut Microbiome Wellness Index 2 (GMWI2) [187]	Determine whether probiotic supplementation improves gut microbiome wellness in healthy individuals and its correlation with clinical markers [197]	Lower baseline GMWI2 identified individuals more responsive to probiotic intervention, while increases in GMWI2 were associated with improved gut microbiome function and lower blood glucose.
Health Index with Principal Component Analysis (hiPCA) [188]	Systematically evaluate and rank dietary factors based on their ability to improve gut microbiota health in overweight individuals [198]	hiPCA values were higher in overweight-associated gut microbiota states and decreased after dietary fermentation, with the largest reductions observed for galacto-oligosaccharides and pueraria polysaccharides.

free exclusive enteral nutrition diet while remaining stable in vegan and omnivorous groups—changes undetectable by α-diversity. In antibiotic-treated cohorts, GMWI2 scores remained depressed for up to six months despite apparent α-diversity recovery; and in fecal microbiota transplantation (FMT) studies, only patients with symptom improvement exhibited increased scores.[187]

Collectively, GMWI2's responsiveness to microbiome dynamics makes it particularly suitable for large-scale dietary research (Figure 2). It facilitates cross-population comparisons of diet-associated wellness, where elevated scores often align with fiber-rich, plant-based diets and lower scores may signal risks from processed, low-fiber foods. Its ability to capture health-relevant shifts missed by α-diversity supports near-real-time monitoring of nutritional interventions. That said, GMWI2 remains indicative rather than diagnostic, and is susceptible to confounders such as geography[199] and recent antibiotic exposure,[200] and—given its taxonomic focus and disease-oriented training—requires additional validation for purely nutritional contexts. A recent transformer-based autoencoder iteration has further improved accuracy and cross-cohort generalizability.[201]

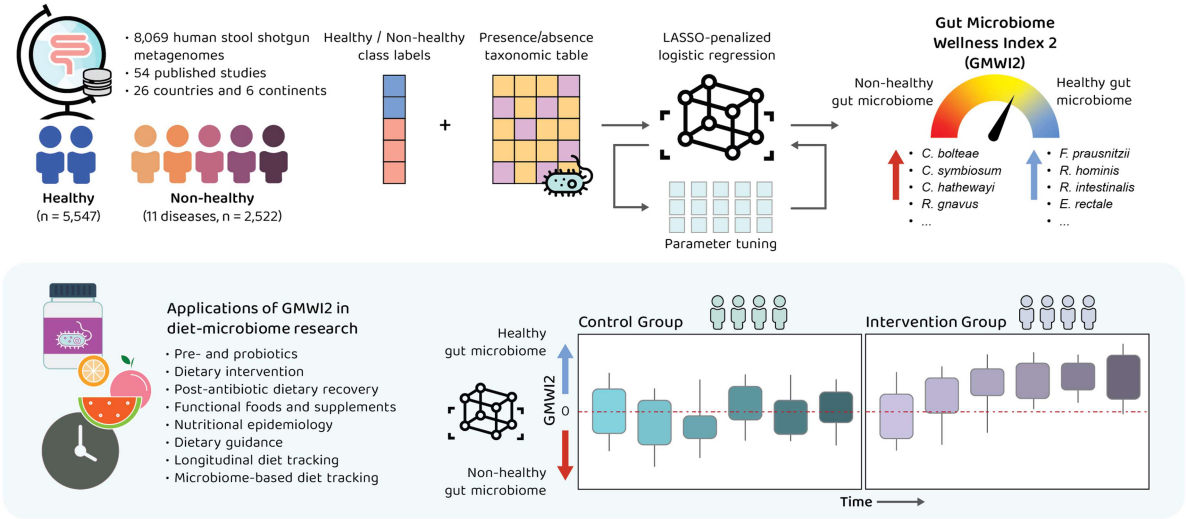

Figure 2. The Gut Microbiome Wellness Index 2 (GMWI2): Construction, interpretation, and applications in diet-gut microbiome research. The upper panel illustrates the GMWI2 workflow: trained on 8,069 human stool shotgun metagenomes from 54 published studies across 26 countries and 6 continents, comprising 5,547 healthy samples and 2,522 samples from 11 disease conditions. Healthy/non-healthy phenotypic labels are combined with a presence/absence taxonomic abundance table, followed by Lasso-penalized logistic regression (with parameter tuning) to select discriminative microbial features. The resulting model outputs a continuous wellness score, where higher values indicate greater relative abundance of health-associated taxa (e.g., *Faecalibacterium prausnitzii*, *Roseburia hominis*, *Roseburia intestinalis*, *Eubacterium rectale*) and lower abundance of non-healthy-associated taxa (e.g., *Clostridium bolteae*, *Clostridium symbiosum*, *Clostridium hathewayi*, *Ruthenibacterium gnavus*). The lower panel highlights GMWI2's utility for monitoring dynamic microbiome responses in nutritional contexts, including pre- and probiotic interventions, dietary trials, post-antibiotic recovery, functional foods/supplements, nutritional epidemiology, personalized dietary guidance, longitudinal diet tracking, and microbiome-based adherence assessment. Time-course examples contrast stable or improving GMWI2 trajectories in intervention arms against declines in controls, demonstrating applicability for detecting diet-induced shifts toward healthier community configurations.

The hiPCA was developed and trained on the same multi-cohort dataset as GMHI.[188] hiPCA uses a principal component-based statistical model to define a "healthy subspace" in metagenomic feature space and quantifies how far individual microbiome profiles deviate from this reference boundary. By evaluating each sample's position relative to this healthy subspace, hiPCA not only identifies deviations indicative of dysbiosis but also pinpoints the specific microbes driving the divergence.

Despite their utility, indices such as GMHI/GMWI2 and hiPCA inherit key limitations from the underlying metagenomic data. Their reliance on species-level taxonomy derived from marker genes precludes strain-level resolution, even though strain variation strongly influences host associations. In addition, these indices capture genomic potential rather than microbial activity and do not reflect dynamic metabolic states or diet-induced functional changes. Consequently, they should be interpreted as associative markers of health-linked microbiome configurations rather than direct measures of microbial function, motivating integration with complementary functional modalities.

While these gut microbiome health metrics are primarily disease-agnostic or condition-specific, a growing focus in nutritional research has led to the development of explicitly diet-specific microbiome metrics. The Dietary Index for Gut Microbiota (DI-GM)[202] stands out, derived from systematic synthesis of over 100 longitudinal human diet-gut microbiome studies. It assigns scores based on intake of fourteen key foods and nutrients: positively weighting fermented dairy (e.g., yogurt), legumes (including chickpeas and soybeans), whole grains, dietary fiber, cranberries, avocados, broccoli, coffee, and green tea, while penalizing refined grains, red and processed meats, and high-fat diets (>40% energy from fat). Unlike general dietary quality scores such as the Healthy Eating Index (HEI) or Mediterranean Diet Score (MDS), the DI-GM was specifically designed to reflect dietary patterns that promote microbial diversity, short-chain fatty acid production, and enrichment of health-associated taxa. Large-cohort validation studies have demonstrated that higher DI-GM scores are associated with lower risks of metabolic syndrome, type 2 diabetes,

cardiovascular disease, IBS, gastrointestinal cancers, and metabolic dysfunction-associated steatotic liver disease, along with reduced inflammation, improved BMI trajectories, and decreased all-cause and cardiovascular mortality. Though relatively new, the DI-GM is a valuable tool for precision nutrition, providing quantitative insights into how daily dietary choices influence gut microbiota composition and function. Its microbiome-specific design complements broader health metrics and holds strong potential for monitoring and optimizing diet-gut microbiome interactions in observational and interventional research.

In contrast to these globally trained, disease-agnostic indices, several population- or condition-specific frameworks have been developed using diverse statistical and machine learning methods. For example, the microbial risk score (MRS)[203] adapts polygenic risk scoring principles to aggregate disease-associated microbial sub-communities into interpretable susceptibility scores, enabling cross-cohort stratification and seamless integration with other omics-derived risk metrics. Another example is the Microbiome Health Index for post-Antibiotic dysbiosis (MHI-A),[204] a simple four-taxon ratio designed to monitor antibiotic-induced disruption and subsequent recovery. Similarly, pediatric microbiota maturity metrics, such as the Microbiota-for-Age Z-score (MAZ),[205] quantify developmental progression of the infant and child gut microbiome relative to chronological age.

These targeted tools are typically trained on specific populations or narrowly defined clinical contexts, prioritizing context-specific precision at the expense of broader generalizability. Nevertheless, many of these indices—or adaptations thereof—hold considerable promise for nutritional research, where they could be used to assess microbiome resilience to dietary perturbations, monitor dynamic responses to nutritional interventions, or delineate diet-associated microbial signatures indicative of gut ecosystem health.

7. Translational priorities and future directions in AI-driven precision nutrition

In this review, we have synthesized the evolving landscape of precision nutrition, emphasizing the central role of diet-gut microbiome interactions in shaping individual metabolic responses. The marked inter-individual variability in dietary response and microbial composition highlights the limitations of traditional one-size-fits-all dietary recommendations and reinforces the need for personalized strategies. Advances in multi-omics profiling (e.g., metagenomics, metatranscriptomics, metaproteomics, and metabolomics) have provided unprecedented resolution into microbial structure, activity, and host-microbiome cross-talk. When integrated with dietary assessments (ranging from self-reported instruments and digital tools to emerging stool-based DNA inference) and host physiological markers (e.g., clinical biomarkers, circadian rhythms, and activity patterns), these modalities create a rich multimodal data ecosystem. Computational workflows, from preprocessing heterogeneous inputs to AI-driven integration, now support applications such as dietary pattern classification, prediction of metabolic phenotypes, responder stratification, and *in silico* simulation of dietary perturbations.

Key innovations include interpretable microbiome health metrics such as GMWI2 and DI-GM, which distill high-dimensional taxonomic data into actionable scores that outperform conventional diversity metrics. These indices enable monitoring of diet-induced shifts and have demonstrated responsiveness in randomized trials, including probiotic interventions associated with improved glycemic control. Complementing data-driven ML approaches, mechanistic GSMMs simulate community-level metabolism and nutrient flux, offering biologically grounded insight into microbe-host interactions. Emerging hybrid frameworks that merge AI-based prediction with mechanistic modeling enhance interpretability, scalability, and generalizability, addressing challenges such as compositionality, high dimensionality, and data heterogeneity.

Despite this progress, fully integrated multi-omics AI systems remain an evolving frontier. More importantly, translating these advances into clinically ready precision nutrition requires deliberate attention to scalability, robustness, and trustworthiness. Several priorities are urgently needed to bridge the gap between research innovation and real-world deployment: First, diverse and longitudinal cohorts remain essential. Many existing predictive models are trained predominantly on Western, adult populations and often rely on dominant modalities such as shotgun metagenomics. Broader inclusion of pediatric, geriatric, multi-ethnic, and chronically ill populations (e.g., diabetes, inflammatory bowel disease, metabolic syndrome) is critical to reduce bias and ensure equitable generalizability. Large-scale initiatives and biobanks integrating microbiome, dietary, and clinical data provide a blueprint, but future efforts must emphasize harmonized protocols, standardized metadata capture, and FAIR-compliant data sharing to enable robust cross-cohort validation.

Second, standardized benchmarking and model evaluation frameworks are urgently required. Current studies frequently use heterogeneous preprocessing pipelines, performance metrics, and validation strategies, limiting reproducibility and cross-study comparability. Shared reference datasets, transparent reporting standards, nested cross-validation, and external validation across independent populations should become the norm. Benchmarking consortia (similar in spirit to CAMI initiatives[206,207]) could be adapted for diet-gut microbiome AI applications, with evaluation criteria that explicitly account for sparsity, compositionality, domain shift, and uncertainty quantification. Third, clinically trustworthy AI systems must incorporate interpretability, uncertainty estimation, and regulatory readiness. Black-box models may achieve high predictive accuracy, but clinical adoption requires transparent reasoning, clear attribution of modifiable dietary drivers, and calibrated uncertainty outputs—particularly in edge cases such as antibiotic-induced dysbiosis or extreme dietary transitions. Hybrid AI-mechanistic models are especially promising in this regard, as they constrain predictions within biologically plausible bounds and provide mechanistic context. Regulatory pathways, including frameworks such as the United States Food and Drug Administration (FDA)'s Software as a Medical Device (SaMD) guidelines, offer potential routes for validation, but early engagement with regulatory and ethical standards is essential. Fourth, scalable clinical integration and closed-loop systems represent the next translational milestone. Real-time feedback architectures—linking wearable sensors (e.g., continuous glucose monitors), digital dietary logging, and periodic microbiome sampling—could enable adaptive dietary recommendations that evolve with an individual's metabolic state.[208] Pragmatic and hybrid clinical trials, including adaptive and N-of-1 designs,[209] are well suited to test these systems under real-world conditions. Partnerships among academia, healthcare systems, and industry will be crucial to generate real-world evidence, integrate tools into electronic health records, and ensure cost-effectiveness beyond specialized research centers. Finally, broader implications must be considered. AI-enhanced precision nutrition holds substantial public health potential, including scalable digital platforms capable of reducing metabolic disease burden. However, overreliance on algorithmic recommendations without behavioral support, equitable access, and bias mitigation could exacerbate disparities. Ethical safeguards, inclusive training datasets, and transparent governance must therefore be embedded from the outset.

Taken together, the coordinated integration of complementary omics technologies, advanced AI, mechanistic modeling, and rigorous translational frameworks is reshaping diet-gut microbiome science. By prioritizing diversity, standardization, interpretability, regulatory alignment, and real-world validation, the field can move beyond proof-of-concept toward clinically actionable precision nutrition. Ultimately, these integrative approaches offer a pathway to transform everyday dietary choices into personalized, evidence-based interventions that promote long-term metabolic health and resilience.

Disclosure of potential conflicts of interest

The authors report there are no competing interests to declare.

Acknowledgments

The study was conceptualized and designed by MAB-S and JS. Data collection and analysis were performed by MAB-S and JS. The original draft of the manuscript was prepared by MAB-S and JS. All authors (MAB-S, CYZ, LVK, and JS) contributed to reviewing and editing the manuscript. All authors approved the final version of the manuscript.

Funding

This study received no external funding.

Author contributions

CRediT: **Maria A. Barrera-Suarez:** Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing; **Conan Y. Zhao:** Writing – review & editing; **Lioudmila V. Karnatovskaia:** Writing – review & editing; **Jaeyun Sung:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Ethical approval

Not applicable.

Clinical trial registration number

Not applicable.

Data availability statement

Not applicable.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. During the preparation of this work, the authors utilized ChatGPT (GPT-5.2, OpenAI) during the final editing stages to improve word choice and readability. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

1. Kolodziejczyk AA, Zheng D, Elinav E. Diet-microbiota interactions and personalized nutrition. Nat Rev Microbiol. 2019;17(12):742–753. doi: 10.1038/s41579-019-0256-8.
2. Berry SE, Valdes AM, Drew DA, Asnicar F, Mazidi M, Wolf J, Capdevila J, Hadjigeorgiou G, Davies R, Al Khatib H, et al. Human postprandial responses to food and potential for precision nutrition. Nat Med. 2020;26(6):964–973. doi: 10.1038/s41591-020-0934-0.
3. Klimenko NS, Odintsova VE, Revel-Muroz A, Tyakht AV. The hallmarks of dietary intervention-resilient gut microbiome. NPJ Biofilms Microbiomes. 2022;8(1):77. doi: 10.1038/s41522-022-00342-8.
4. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL, et al. The long-term stability of the human gut microbiota. Science. 2013;341(6141):1237439. doi: 10.1126/science.1237439.
5. Ross FC, Patangia D, Grimaud G, Lavelle A, Dempsey EM, Ross RP, Stanton C. The interplay between diet and the gut microbiome: implications for health and disease. Nat Rev Microbiol. 2024;22(11):671–686. doi: 10.1038/s41579-024-01068-4.
6. Zhang P. Influence of foods and nutrition on the gut microbiome and implications for intestinal health. Int J Mol Sci. 2022;23(17):9588. doi: 10.3390/ijms23179588.
7. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, Knight R, Manjurano A, Changalucha J, Elias JE, et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. Science. 2017;357(6353):802–806. doi: 10.1126/science.aan4834.
8. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni S, Biagi E, Peano C, Severgnini M, et al. Gut microbiome of the hadza hunter-gatherers. Nat Commun. 2014;5(1):3654. doi: 10.1038/ncomms4654.
9. Vangay P, Johnson AJ, Ward TL, Al-Ghalith GA, Shields-Cutler RR, Hillmann BM, Lucas SK, Beura LK, Thompson EA, Till LM, et al. US immigration westernizes the human gut microbiome. Cell. 2018;175(4):962–972.e10. doi: 10.1016/j.cell.2018.10.029.
10. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444(7122):1027–1031.
11. Beam A, Clinger E, Hao L. Effect of diet and dietary components on the composition of the gut microbiota. Nutrients. 2021;13(8):2795. doi: 10.3390/nu13082795.
12. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2014;505(7484):559–563. doi: 10.1038/nature12820.
13. Bourdeau-Julien I, Castonguay-Paradis S, Rochefort G, Perron J, Lamarche B, Flamand N, Di Marzo V, Veilleux A, Raymond F. The diet rapidly and differentially affects the gut microbiota and host lipid mediators in a healthy population. Microbiome. 2023;11(1):26. doi: 10.1186/s40168-023-01469-2.
14. Pascale A, Marchesi N, Marelli C, Coppola A, Luzi L, Govoni S, Giustina A, Gazzaruso C. Microbiota and metabolic diseases. Endocrine. 2018;61(3):357–371. doi: 10.1007/s12020-018-1605-5.
15. Jiang W, Wu N, Wang X, Chi Y, Zhang Y, Qiu X, Hu Y, Li J, Liu Y. Dysbiosis gut microbiota associated with inflammation and impaired mucosal immune function in intestine of humans with non-alcoholic fatty liver disease. Sci Rep. 2015;5(1):8096. doi: 10.1038/srep08096.

16. Chelakkot C, Ghim J, Ryu SH. Mechanisms regulating intestinal barrier integrity and its pathological implications. Exp Mol Med. 2018;50(8):1–9. doi: 10.1038/s12276-018-0126-x.
17. Heiss CN, Olofsson LE. Gut microbiota-dependent modulation of energy metabolism. J Innate Immun. 2018;10(3):163–171. doi: 10.1159/000481519.
18. Singh RK, Chang H-W, Yan D, Lee KM, Ucmak D, Wong K, Abrouk M, Farahnik B, Nakamura M, Zhu TH, et al. Influence of diet on the gut microbiome and implications for human health. J Transl Med. 2017;15(1):73. doi: 10.1186/s12967-017-1175-y.
19. Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, Semenkovich CF, Gordon JI. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A. 2004;101(44):15718–15723. doi: 10.1073/pnas.0407076101.
20. Tilg H, Moschen AR. Microbiota and diabetes: an evolving relationship. Gut. 2014;63(9):1513–1521. doi: 10.1136/gutjnl-2014-306928.
21. Zhuang Z-Q, Shen L-L, Li W-W, Fu X, Zeng F, Gui L, Lü Y, Cai M, Zhu C, Tan Y-L, et al. Gut microbiota is altered in patients with Alzheimer's disease. J Alzheimers Dis. 2018;63(4):1337–1346. doi: 10.3233/JAD-180176.
22. Sherwin E, Dinan TG, Cryan JF. Recent developments in understanding the role of the gut microbiota in brain health and disease. Ann N Y Acad Sci. 2018;1420(1):5–25. doi: 10.1111/nyas.13416.
23. Manichanh C, Borruel N, Casellas F, Guarner F. The gut microbiota in IBD. Nat Rev Gastroenterol Hepatol. 2012;9(10):599–608. doi: 10.1038/nrgastro.2012.152.
24. Ordovas JM, Ferguson LR, Tai ES, Mathers JC. Personalised nutrition and health. BMJ. 2018;361 bmj.k2173. doi: 10.1136/bmj.k2173.
25. Marcos-Zambrano LJ, Karaduzovic-Hadziabdic K, Loncar Turukalo T, Przymus P, Trajkovik V, Aasmets O, Berland M, Gruca A, Hasic J, Hron K, et al. Applications of machine learning in human microbiome studies: a review on feature selection, biomarker identification, disease prediction and treatment. Front Microbiol. 2021;12:634511. doi: 10.3389/fmicb.2021.634511.
26. Theodore Armand TP, Nfor KA, Kim J-I, Kim H-C. Applications of artificial intelligence, machine learning, and deep learning in nutrition: a systematic review. Nutrients. 2024;16(7):1073. doi: 10.3390/nu16071073.
27. Liu Y-Y. Deep learning for microbiome-informed precision nutrition. Natl Sci Rev. 2025;12(6):nwaf148. doi: 10.1093/nsr/nwaf148.
28. Wang X, Sun Z, Xue H, An R. Artificial intelligence applications to personalized dietary recommendations: a systematic review. Healthcare (Basel). 2025;13(12):1417. doi: 10.3390/healthcare13121417.
29. Kumar B, Lorusso E, Fosso B, Pesole G. A comprehensive overview of microbiome data in the light of machine learning applications: categorization, accessibility, and future directions. Front Microbiol. 2024;15:1343572. doi: 10.3389/fmicb.2024.1343572.
30. 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.
31. Mallick H, Ma S, Franzosa EA, Vatanen T, Morgan XC, Huttenhower C. Experimental design and quantitative analysis of microbial community multiomics. Genome Biol. 2017;18(1):228. doi: 10.1186/s13059-017-1359-z.
32. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Comput Sci. 2017;3:e104e104. doi: 10.7717/peerj-cs.104.
33. Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, Manghi P, Dubois L, Huang KD, Thomas AM, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol. 2023;41(11):1633–1644. doi: 10.1038/s41587-023-01688-w.
34. Kim Y, Worby CJ, Acharya S, van Dijk LR, Alfonsetti D, Gromko Z, Azimzadeh PN, Dodson KW, Gerber GK, Hultgren SJ, et al. Longitudinal profiling of low-abundance strains in microbiomes with ChronoStrain. Nat Microbiol. 2025;10(5):1184–1197. doi: 10.1038/s41564-025-01983-z.
35. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Qin J, et al., MetaHIT Consortium A human gut microbial gene catalogue established by metagenomic sequencing. Nature. 2010;464(7285):59–65. doi: 10.1038/nature08821.
36. Commichaux S, Shah N, Ghurye J, Stoppel A, Goodheart JA, Luque GG, Cummings MP, Pop M. A critical assessment of gene catalogs for metagenomic analysis. Bioinformatics. 2021;37(18):2848–2857. doi: 10.1093/bioinformatics/btab216.
37. Bishara A, Moss EL, Kolmogorov M, Parada AE, Weng Z, Sidow A, Dekas AE, Batzoglou S, Bhatt AS. High-quality genome sequences of uncultured microbes by assembly of read clouds. Nat Biotechnol. 2018;36(11):1067–1075. doi: 10.1038/nbt.4266.
38. Benoit G, Raguideau S, James R, Phillippy AM, Chikhi R, Quince C. High-quality metagenome assembly from long accurate reads with metaMDBG. Nat Biotechnol. 2024;42(9):1378–1383. doi: 10.1038/s41587-023-01983-6.
39. Bharti R, Grimm DG. Current challenges and best-practice protocols for microbiome analysis. Brief Bioinform. 2021;22(1):178–193. doi: 10.1093/bib/bbz155.
40. Regueira-Iglesias A, Balsa-Castro C, Blanco-Pintos T, Tomás I. Critical review of 16S rRNA gene sequencing workflow in microbiome studies: from primer selection to advanced data analysis. Mol Oral Microbiol. 2023;38(5):347–399. doi: 10.1111/omi.12434.

41. Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchel T, Perry T, Kao D, Mason AL, Madsen KL, et al. Characterization of the gut microbiome using 16S or shotgun metagenomics. Front Microbiol. 2016;7:459. doi: 10.3389/fmicb.2016.00459.
42. Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, List M, Neuhaus K. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. mSphere [Internet]. 2021;6(1):e01202-20. [accessed 2025 Apr 25]. doi: 10.1128/mSphere.01202-20 .
43. Wemheuer F, Taylor JA, Daniel R, Johnston E, Meinicke P, Thomas T, Wemheuer B. Tax4Fun2: prediction of habitat-specific functional profiles and functional redundancy based on 16S rRNA gene sequences. Environ Microbiome. 2020;15(1):11. doi: 10.1186/s40793-020-00358-7.
44. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol. 2020;38(6):685–688. doi: 10.1038/s41587-020-0548-6.
45. Pereira-Marques J, Hout A, Ferreira RM, Weber M, Pinto-Ribeiro I, van Doorn L-J, Knetsch CW, Figueiredo C. Impact of host DNA and sequencing depth on the taxonomic resolution of whole metagenome sequencing for microbiome analysis. Front Microbiol. 2019;10:1277. doi: 10.3389/fmicb.2019.01277.
46. McLaren MR, Willis AD, Callahan BJ. Consistent and correctable bias in metagenomic sequencing experiments. eLife. 2019;8:e46923. doi: 10.7554/eLife.46923.
47. Chiu CY, Miller SA. Clinical metagenomics. Nat Rev Genet. 2019;20(6):341–355. doi: 10.1038/s41576-019-0113-7.
48. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. Front Microbiol. 2017;8:2224. doi: 10.3389/fmicb.2017.02224.
49. Shakya M, Lo C-C, Chain PSG. Advances and challenges in metatranscriptomic analysis. Front Genet. 2019;10:904. doi: 10.3389/fgene.2019.00904.
50. Deutscher MP. Degradation of RNA in bacteria: comparison of mRNA and stable RNA. Nucleic Acids Res. 2006;34(2):659–666. doi: 10.1093/nar/gkj472.
51. Ojala T, Häkkinen A-E, Kankuri E, Kankainen M. Current concepts, advances, and challenges in deciphering the human microbiota with metatranscriptomics. Trends Genet. 2023;39(9):686–702. doi: 10.1016/j.tig.2023.05.004.
52. Xiong W, Abraham PE, Li Z, Pan C, Hettich RL. Microbial metaproteomics for characterizing the range of metabolic functions and activities of human gut microbiota. Proteomics. 2015;15(20):3424–3438. doi: 10.1002/pmic.201400571.
53. Kleiner M. Metaproteomics: much more than measuring gene expression in microbial communities. mSystems. 2019;4(3):e00115-19. doi: 10.1128/mSystems.00115-19.
54. Hettich RL, Pan C, Chourey K, Giannone RJ. Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. Anal Chem. 2013;85(9):4203–4214. doi: 10.1021/ac303053e.
55. Xian F, Brenek M, Krisp C, Urbauer E, Ravi Kumar RK, Aguanno D, Srikumar T, Liu Q, Barry AM, Ma B, et al. Ultra-sensitive metaproteomics redefines the dark metaproteome, uncovering host-microbiome interactions and drug targets in intestinal diseases. Nat Commun. 2025;16(1):6644. doi: 10.1038/s41467-025-61977-7.
56. Tanca A, Palomba A, Fiorito G, Abbondio M, Pagnozzi D, Uzzau S. Metaproteomic portrait of the healthy human gut microbiota. NPJ Biofilms Microbiomes. 2024;10(1):54. doi: 10.1038/s41522-024-00526-4.
57. Lee PY, Chin S-F, Neoh H-M, Jamal R. Metaproteomic analysis of human gut microbiota: where are we heading? J Biomed Sci. 2017;24(1):36. doi: 10.1186/s12929-017-0342-z.
58. Krautkramer KA, Fan J, Bäckhed F. Gut microbial metabolites as multi-kingdom intermediates. Nat Rev Microbiol. 2021;19(2):77–94. doi: 10.1038/s41579-020-0438-4.
59. Fu Y, Lyu J, Wang S. The role of intestinal microbes on intestinal barrier function and host immunity from a metabolite perspective. Front Immunol. 2023;14:1277102. doi: 10.3389/fimmu.2023.1277102.
60. Ghosh S, Whitley CS, Haribabu B, Jala VR. Regulation of intestinal barrier function by microbial metabolites. Cell Mol Gastroenterol Hepatol. 2021;11(5):1463–1482. doi: 10.1016/j.jcmgh.2021.02.007.
61. Jyoti, Dey P. Mechanisms and implications of the gut microbial modulation of intestinal metabolic processes. NPJ Metab Health Dis. 2025;3(1):24. doi: 10.1038/s44324-025-00066-1.
62. Moya A, Ferrer M. Functional redundancy-induced stability of gut microbiota subjected to disturbance. Trends Microbiol. 2016;24(5):402–413. doi: 10.1016/j.tim.2016.02.002.
63. Comte J, Fauteux L, Del Giorgio PA. Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. Front Microbiol. 2013;4:112. doi: 10.3389/fmicb.2013.00112.
64. Allison SD, Martiny JBH. Colloquium paper: resistance, resilience, and redundancy in microbial communities. Proc Natl Acad Sci U S A. 2008;105 Suppl 1(supplement_1):11512–11519. doi: 10.1073/pnas.0801925105.
65. O'Connell TM. The application of metabolomics to probiotic and prebiotic interventions in human clinical studies. Metabolites. 2020;10(3):120. doi: 10.3390/metabo10030120.
66. Yang Q, Zhang A-H, Miao J-H, Sun H, Han Y, Yan G-L, Wu F-F, Wang X-J. Metabolomics biotechnology, applications, and future trends: a systematic review. RSC Adv. 2019;9(64):37245–37257. doi: 10.1039/C9RA06697G.

67. Smirnov KS, Maier TV, Walker A, Heinzmann SS, Forcisi S, Martinez I, Walter J, Schmitt-Kopplin P. Challenges of metabolomics in human gut microbiota research. Int J Med Microbiol. 2016;306(5):266–279. doi: 10.1016/j.ijmm.2016.03.006.
68. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 2016;3(1):160018. doi: 10.1038/sdata.2016.18.
69. Shim J-S, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. Epidemiol Health. 2014;36:e2014009. doi: 10.4178/epih/e2014009.
70. Kirkpatrick SI, Subar AF, Douglass D, Zimmerman TP, Thompson FE, Kahle LL, George SM, Dodd KW, Potischman N. Performance of the automated self-administered 24-hour recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. AJCN. 2014;100(1):233–240. doi: 10.3945/ajcn.114.083238.
71. U.S. Department of Agriculture, Agricultural Research Service. (2019) USDA FoodData Central. fdc.nal.usda.gov [Internet]. [accessed 2025 Dec 14]. https://fdc.nal.usda.gov/
72. Health Canada. Canadian Nutrient File (CNF) [Internet]. [accessed 2025 Dec 18]. 2024. https://food-nutrition.canada.ca/cnf-fce/
73. Barabási A-L, Menichetti G, Loscalzo J. The unmapped chemical complexity of our diet. Nat Food. 2019;1(1):33–37. doi: 10.1038/s43016-019-0005-1.
74. Mezgec S, Koroušić Seljak B. NutriNet: a deep learning food and drink image recognition system for dietary assessment. Nutrients. 2017;9(7):657. doi: 10.3390/nu9070657.
75. Lu Y, Stathopoulou T, Vasiloglou MF, Pinault LF, Kiley C, Spanakis EK, Mougiakakou S. GoFOODTM: an artificial intelligence system for dietary assessment. Sensors (Basel). 2020;20(15):4283. doi: 10.3390/s20154283.
76. Ahmad Z, Kerr DA, Bosch M, Boushey CJ, Delp EJ, Khanna N, Zhu F. A mobile food record for integrated dietary assessment. MADiMa16 (2016). 2016;16:53–62. doi: 10.1145/2986035.2986038.
77. Myers A, Johnston N, Rathod V, Korattikara A, Gorban A, Silberman N, Guadarrama S, Papandreou G, Huang J, Murphy K. Im2Calories: towards an automated mobile vision food diary. In 2015 IEEE International Conference on Computer Vision (ICCV). 2015. p. 1233–1241. IEEE. doi: 10.1109/ICCV.2015.146.
78. Joutou T, Yanai K. A food image recognition system with multiple kernel learning. In 2009 16th IEEE International Conference on Image Processing (ICIP). 2009. Vol. 285–288, p. 285–288. IEEE. doi: 10.1109/ICIP.2009.5413400.
79. Lamarine M, Hager J, Saris WHM, Astrup A, Valsesia A. Fast and accurate approaches for large-scale, automated mapping of food diaries on food composition tables. Front Nutr. 2018;5:38. doi: 10.3389/fnut.2018.00038.
80. Bingham SA. Urine nitrogen as a biomarker for the validation of dietary protein intake. J Nutr. 2003;133 Suppl 3(3):921S–924S. doi: 10.1093/jn/133.3.921S.
81. Radtke MD, Poe M, Stookey J, Jilcott Pitts S, Moran NE, Landry MJ, Rubin LP, Stage VC, Scherr RE. Recommendations for the use of the veggie Meter® for spectroscopy-based skin carotenoid measurements in the research setting. Curr Dev Nutr. 2021;5(8):nzab104. doi: 10.1093/cdn/nzab104.
82. Petzke KJ, Boeing H, Klaus S, Metges CC. Carbon and nitrogen stable isotopic composition of hair protein and amino acids can be used as biomarkers for animal-derived dietary protein intake in humans. J Nutr. 2005;135(6):1515–1520. doi: 10.1093/jn/135.6.1515.
83. Nash SH, Kristal AR, Bersamin A, Hopkins SE, Boyer BB, O'Brien DM. Carbon and nitrogen stable isotope ratios predict intake of sweeteners in a yup'ik study population. J Nutr. 2013;143(2):161–165. doi: 10.3945/jn.112.169425.
84. Scalbert A, Brennan L, Manach C, Andres-Lacueva C, Dragsted LO, Draper J, Wishart D. The food metabolome: a window over dietary exposure. Am J Clin Nutr. 2014;99(6):1286–1308. doi: 10.3945/ajcn.113.076133.
85. Mennen LI, Sapinho D, Ito H, Bertrais S, Galan P, Hercberg S, Scalbert A. Urinary flavonoids and phenolic acids as biomarkers of intake for polyphenol-rich foods. Br J Nutr. 2006;96(1):191–198. doi: 10.1079/BJN20061808.
86. Pope R, Visconti A, Zhang X, Louca P, Baleanu A-F, Lin Y, Asnicar F, Bermingham K, Wong KE, Michelotti GA, et al. Faecal metabolites as a readout of habitual diet capture dietary interactions with the gut microbiome. Nat Commun. 2025;16(1):10051. doi: 10.1038/s41467-025-66046-7.
87. Valdés-Mas R, Leshem A, Zheng D, Cohen Y, Kern L, Zmora N, He Y, Katina C, Eliyahu-Miller S, Yosef-Hevroni T, et al. Metagenome-informed metaproteomics of the human gut microbiome, host, and dietary exposome uncovers signatures of health and inflammatory bowel disease. Cell. 2025;188(4):1062–1083.e36. doi: 10.1016/j.cell.2024.12.016.
88. Superdock DK, Petrone BL, Kirtley MC, David LA. From stool to sequence: decoding the human diet with FoodSeq. mSystems. 2025;10(7):e0015825. doi: 10.1128/msystems.00158-25.
89. Diener C, Holscher HD, Filek K, Corbin KD, Moissl-Eichinger C, Gibbons SM. Metagenomic estimation of dietary intake from human stool. Nat Metab. 2025;7(3):617–630. doi: 10.1038/s42255-025-01220-1.
90. Aqeel A, Kay MC, Zeng J, Petrone BL, Yang C, Truong T, Brown CB, Jiang S, Carrion VM, Bryant S, et al. Grocery intervention and DNA-based assessment to improve diet quality in pediatric obesity: a pilot randomized controlled study. Obesity (Silver Spring). 2025;33(2):331–345. doi: 10.1002/oby.24205.

91. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, Ben-Yacov O, Lador D, Avnit-Sagi T, Lotan-Pompan M, et al. Personalized nutrition by prediction of glycemic responses. Cell. 2015;163(5):1079–1094. doi: 10.1016/j.cell.2015.11.001.
92. Joshi S, Kesavadev J, K M PK, Saboo B, Mehta A, Bhattacharyya A, Sosale A, Jabbar PK, Santosh R, Deshmukh V, et al. Postprandial glucose: a variable in continuum. Clin Med Insights Endocrinol Diabetes. 2025;18(11795514251370508):11795514251370508. doi: 10.1177/11795514251370507.
93. Sung J, Kim S, Cabatbat JJT, Jang S, Jin Y-S, Jung GY, Chia N, Kim P-J. Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. Nat Commun. 2017;8:15393. doi: 10.1038/ncomms15393.
94. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–444.
95. Schmidhuber J (2014) Deep learning in neural networks: An overview. arXiv [csNE] [Internet] http://arxiv.org/abs/1404.7828.
96. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature. 1986;323(6088):533–536. doi: 10.1038/323533a0.
97. Hernández Medina R, Kutuzova S, Nielsen KN, Johansen J, Hansen LH, Nielsen M, Rasmussen S. Machine learning and deep learning applications in microbiome research. ISME Commun. 2022;2(1):98. doi: 10.1038/s43705-022-00182-9.
98. Przymus P, Rykaczewski K, Martín-Segura A, Truu J, Carrillo De Santa Pau E, Kolev M, Naskinova I, Gruca A, Sampri A, Frohme M, et al. Deep learning in microbiome analysis: a comprehensive review of neural network models. Front Microbiol. 2024;15:1516667. doi: 10.3389/fmicb.2024.1516667.
99. Fackelmann G, Manghi P, Carlino N, Heidrich V, Piccinno G, Ricci L, Piperni E, Arrè A, Bakker E, Creedon AC, et al. Gut microbiome signatures of vegan, vegetarian and omnivore diets and associated health outcomes across 21,561 individuals. Nat Microbiol. 2025;10(1):41–52. doi: 10.1038/s41564-024-01870-z.
100. Tarallo S, Ferrero G, De Filippis F, Francavilla A, Pasolli E, Panero V, Cordero F, Segata N, Grioni S, Pensa RG, et al. Stool microRNA profiles reflect different dietary and gut microbiome patterns in healthy individuals. Gut. 2022;71(7):1302–1314. doi: 10.1136/gutjnl-2021-325168.
101. Singh R, McDonald D, Hernandez AR, Song SJ, Bartko A, Knight R, Salathé M. Temporal nutrition analysis associates dietary regularity and quality with gut microbiome diversity: insights from the food & you digital cohort. Nat Commun. 2025;16(1):8635. doi: 10.1038/s41467-025-63799-z.
102. Bisanz JE, Upadhyay V, Turnbaugh JA, Ly K, Turnbaugh PJ. Meta-analysis reveals reproducible gut microbiome alterations in response to a high-fat diet. Cell Host Microbe. 2019;26(2):265–272.e4. doi: 10.1016/j.chom.2019.06.013.
103. Mendes-Soares H, Raveh-Sadka T, Azulay S, Ben-Shlomo Y, Cohen Y, Ofek T, Stevens J, Bachrach D, Kashyap P, Segal L, et al. Model of personalized postprandial glycemic response to food developed for an Israeli cohort predicts responses in midwestern American individuals. AJCN. 2019;110(1):63–75. doi: 10.1093/ajcn/nqz028.
104. Ben-Yacov O, Godneva A, Rein M, Shilo S, Lotan-Pompan M, Weinberger A, Segal E. Gut microbiome modulates the effects of a personalised postprandial-targeting (PPT) diet on cardiometabolic markers: a diet intervention in pre-diabetes. Gut. 2023;72(8):1486–1496. doi: 10.1136/gutjnl-2022-329201.
105. Rein M, Ben-Yacov O, Godneva A, Shilo S, Zmora N, Kolobkov D, Cohen-Dolev N, Wolf B-C, Kosower N, Lotan-Pompan M, et al. Effects of personalized diets by prediction of glycemic responses on glycemic control and metabolic health in newly diagnosed T2DM: a randomized dietary intervention pilot trial. BMC Med. 2022;20(1):56. doi: 10.1186/s12916-022-02254-y.
106. Popova PV, Isakov AO, Rusanova AN, Sitkin SI, Anopova AD, Vasukova EA, Tkachuk AS, Nemikina IS, Stepanova EA, Eriskovskaya AI, et al. Personalized prediction of glycemic responses to food in women with diet-treated gestational diabetes: the role of the gut microbiota. NPJ Biofilms Microbiomes. 2025;11(1):25. doi: 10.1038/s41522-025-00650-9.
107. Asnicar F, Berry SE, Valdes AM, Nguyen LH, Piccinno G, Drew DA, Leeming E, Gibson R, Le Roy C, Khatib HA, et al. Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. Nat Med. 2021;27(2):321–332. doi: 10.1038/s41591-020-01183-8.
108. Jeong K, Moon S-J, Rachim VP, Song Y, Cho Y-M, Park S-M. Enhanced post-prandial glycemic response prediction in type 2 diabetes with microbiome data and deep learning. IEEE J Biomed Health Inform. 2026;30(1):643–654.
109. Shi N, Nepal S, Hoobler R, Menni C, Playdon MC, Spakowicz D, Wells PM, Steves CJ, Clinton SK, Tabung FK. Pro-inflammatory and hyperinsulinaemic dietary patterns are associated with specific gut microbiome profiles: a TwinsUK cohort study. Gut Microbiome (Camb). 2024;5:e12. doi: 10.1017/gmb.2024.14.
110. Bolte LA, Vich Vila A, Imhann F, Collij V, Gacesa R, Peters V, Wijmenga C, Kurilshikov A, Campmans-Kuijpers MJE, Fu J, et al. Long-term dietary patterns are associated with pro-inflammatory and anti-inflammatory features of the gut microbiome. Gut. 2021;70(7):1287–1298. doi: 10.1136/gutjnl-2020-322670.
111. Tang Z-Z, Chen G, Hong Q, Huang S, Smith HM, Shah RD, Scholz M, Ferguson JF. Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. Front Genet. 2019;10:454. doi: 10.3389/fgene.2019.00454.

112. Shinn LM, Mansharamani A, Baer DJ, Novotny JA, Charron CS, Khan NA, Zhu R, Holscher HD. Fecal metabolites as biomarkers for predicting food intake by healthy adults. J Nutr. 2023;152(12):2956–2965. doi: 10.1093/jn/nxac195.
113. Prendiville O, Walton J, Flynn A, Nugent AP, McNulty BA, Brennan L. Classifying individuals into a dietary pattern based on metabolomic data. Mol Nutr Food Res. 2021;65(11):e2001183. doi: 10.1002/mnfr.202001183.
114. Baldeon AD, Holthaus TA, Khan NA, Holscher HD. Fecal microbiota and metabolites predict metabolic health features across various dietary patterns in adults. J Nutr. 2025;155(6):1795–1803. doi: 10.1016/j.tjnut.2025.03.024.
115. Tily H, Patridge E, Cai Y, Gopu V, Gline S, Genkin M, Lindau H, Sjue A, Slavov I, Perlina A, et al. Gut microbiome activity contributes to prediction of individual variation in glycemic response in adults. Diabetes Ther. 2022;13(1):89–111. doi: 10.1007/s13300-021-01174-z.
116. Popp CJ, Hu L, Kharmats AY, Curran M, Berube L, Wang C, Pompeii ML, Illiano P, St-Jules DE, Mottern M, et al. Effect of a personalized diet to reduce postprandial glycemic response vs a low-fat diet on weight loss in adults with abnormal glucose metabolism and obesity: a randomized clinical trial. JAMA Netw Open. 2022;5(9):e2233760. doi: 10.1001/jamanetworkopen.2022.33760.
117. Alherfi MM, Khazna MS, Aldakhil AM, Alghamdi SS, Altahhan ME, Alalwani JF, Asiri FSH, Alkhars AHA, Mohaini MA. Precision modulation of gut microbiome via metagenomic-driven dietary and probiotic interventions for non-alcoholic fatty liver disease management in obese Saudis. J Adv Trends Med Res. 2025;2(3):596–601. doi: 10.4103/ATMR.ATMR_101_25.
118. Bermingham KM, Linenberg I, Polidori L, Asnicar F, Arrè A, Wolf J, Badri F, Bernard H, Capdevila J, Bulsiewicz WJ, et al. Effects of a personalized nutrition program on cardiometabolic health: a randomized controlled trial. Nat Med. 2024;30(7):1888–1897. doi: 10.1038/s41591-024-02951-6.
119. Connell J, Toma R, Ho CH-C, Shen N, Moura P, Le T, Patridge E, Antoine G, Keiser H, Julian C, et al. Data-driven precision nutrition improves clinical outcomes and risk scores for IBS, depression, anxiety, and T2D. Am J Lifestyle Med [Internet]. 2023;15598276231216392. [accessed 2026 Feb 9]. doi: 10.1177/15598276231216393.
120. Kallapura G, Prakash AS, Sankaran K, Manjappa P, Chaudhary P, Ambhore S, Dhar D. Microbiota based personalized nutrition improves hyperglycaemia and hypertension parameters and reduces inflammation: a prospective, open label, controlled, randomized, comparative, proof of concept study. PeerJ. 2024;12:e17583e17583. doi: 10.7717/peerj.17583.
121. Tunali V, Arslan N, Ermiş BH, Derviş Hakim G, Gündoğdu A, Hora M, Nalbantoğlu [. A multicenter randomized controlled trial of microbiome-based artificial intelligence-assisted personalized diet vs low fodmap diet: a novel approach for the management of irritable bowel syndrome. Am J Gastroenterol. 2024;119(9):1901–1912. doi: 10.14309/ajg.0000000000002862.
122. Nielsen RL, Helenius M, Garcia SL, Roager HM, Aytan-Aktug D, Hansen LBS, Lind MV, Vogt JK, Dalgaard MD, Bahl MI, et al. Data integration for prediction of weight loss in randomized controlled dietary trials. Sci Rep. 2020;10(1):20103. doi: 10.1038/s41598-020-76097-z.
123. Hoffmann Sardá FA, Giuntini EB, Oliveira A, Souza GS, Prado SBR, Taddei CR, Tadini CC, Bittinger K, Bushman FD, Menezes EW, et al. Baseline intestinal microbiota composition influences response to a real-world dietary fiber intervention. NPJ Biofilms Microbiomes. 2025;11(1):203. doi: 10.1038/s41522-025-00817-4.
124. Dong TS, Luu K, Lagishetty V, Sedighian F, Woo S-L, Dreskin BW, Katzka W, Chang C, Zhou Y, Arias-Jayo N, et al. The intestinal microbiome predicts weight loss on a calorie-restricted diet and is associated with improved hepatic steatosis. Front Nutr. 2021;8:718661. doi: 10.3389/fnut.2021.718661.
125. Huang C, Liu D, Yang S, Huang Y, Wei X, Zhang P, Lin J, Xu B, Liu Y, Guo D, et al. Effect of time-restricted eating regimen on weight loss is mediated by gut microbiome. iSci. 2024;27(7):110202. doi: 10.1016/j.isci.2024.110202.
126. Klimenko NS, Tyakht AV, Popenko AS, Vasiliev AS, Altukhov IA, Ischenko DS, Shashkova TI, Efimova DA, Nikogosov DA, Osipenko DA, et al. Microbiome responses to an uncontrolled short-term diet intervention in the frame of the citizen science project. Nutrients. 2018;10(5):576. doi: 10.3390/nu10050576.
127. Song D, Feng G, Ma Y, Shi Y, Qian C, Wang C, Xu J, Li Y, Wang X, Fan N, et al. Gut microbiome predicts personalized responses to dietary fiber in prediabetes: a randomized, open-label trial. Nat Commun. 2025;16(1):11506. doi: 10.1038/s41467-025-66498-x.
128. Kouraki A, Nogal A, Nocun W, Louca P, Vijay A, Wong K, Michelotti GA, Menni C, Valdes AM. Machine learning metabolomics profiling of dietary interventions from a six-week randomised trial. Metabolites. 2024;14(6):311. doi: 10.3390/metabo14060311.
129. Cheng R, Wang L, Le S, Yang Y, Zhao C, Zhang X, Yang X, Xu T, Xu L, Wiklund P, et al. A randomized controlled trial for response of microbiome network to exercise and diet intervention in patients with nonalcoholic fatty liver disease. Nat Commun. 2022;13(1):2555. doi: 10.1038/s41467-022-29968-0.
130. Wang T, Holscher HD, Maslov S, Hu FB, Weiss ST, Liu Y-Y. Predicting metabolite response to dietary intervention using deep learning. Nat Commun. 2025;16(1):815. doi: 10.1038/s41467-025-56165-6.
131. Reiman D, Dai Y. Using autoencoders for predicting latent microbiome community shifts responding to dietary changes. In 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2019. p. 1884–1891. IEEE. doi: 10.1109/BIBM47256.2019.8983124.

132. 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;18:100607100607. doi: 10.1016/j.mlwa.2024.100607.
133. Räz T. ML interpretability: simple isn't easy. Stud Hist Philos Sci. 2024;103:159–167. doi: 10.1016/j.shpsa.2023.12.007.
134. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics. 2010;26(10):1340–1347. doi: 10.1093/bioinformatics/btq134.
135. Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet]. 2016. p. 1135–1144. New York, NY, USA: ACM. doi: 10.1145/2939672.2939778.
136. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. arXiv [csAI] [Internet]. [accessed 2025 Dec 19]. 2017. doi: 10.48550/arXiv.1705.07874
137. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: explaining supervised machine learning model predictions in drug development. Clin Transl Sci. 2024;17(11):e70056. doi: 10.1111/cts.70056.
138. Szczepankiewicz K, Popowicz A, Charkiewicz K, Nałęcz-Charkiewicz K, Szczepankiewicz M, Lasota S, Zawistowski P, Radlak K. Ground truth based comparison of saliency maps algorithms. Sci Rep. 2023;13(1):16887. doi: 10.1038/s41598-023-42946-w.
139. Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. Neurocomputing. 2021;452:48–62. doi: 10.1016/j.neucom.2021.03.091.
140. Guidotti R, Monreale A, Ruggieri S, Turini F, Giannotti F, Pedreschi D. A survey of methods for explaining black box models. ACM Comput Surv. 2019;51(5):1–42. doi: 10.1145/3236009.
141. Lee BY, Ordovás JM, Parks EJ, Anderson CAM, Barabási A-L, Clinton SK, de la Haye K, Duffy VB, Franks PW, Ginexi EM, et al. Research gaps and opportunities in precision nutrition: an NIH workshop report. AJCN. 2022;116(6):1877–1900. doi: 10.1093/ajcn/nqac237.
142. Zeisel SH. A conceptual framework for studying and investing in precision nutrition. Front Genet. 2019;10:200. doi: 10.3389/fgene.2019.00200.
143. Guizar-Heredia R, Noriega LG, Rivera AL, Resendis-Antonio O, Guevara-Cruz M, Torres N, Tovar AR. A new approach to personalized nutrition: postprandial glycemic response and its relationship to gut microbiota. Arch Med Res. 2023;54(3):176–188. doi: 10.1016/j.arcmed.2023.02.007.
144. Ben-Yacov O, Godneva A, Rein M, Shilo S, Kolobkov D, Koren N, Cohen Dolev N, Travinsky Shmul T, Wolf BC, Kosower N, et al. Personalized postprandial glucose response-targeting diet versus Mediterranean diet for glycemic control in prediabetes. Diabetes Care. 2021;44(9):1980–1991. doi: 10.2337/dc21-0162.
145. Karakan T, Gundogdu A, Alagözlü H, Ekmen N, Ozgul S, Tunali V, Hora M, Beyazgul D, Nalbantoglu OU. Artificial intelligence-based personalized diet: a pilot clinical study for irritable bowel syndrome. Gut Microbes. 2022;14(1):2138672. doi: 10.1080/19490976.2022.2138672.
146. Mallick H, Franzosa EA, Mclver LJ, Banerjee S, Sirota-Madi A, Kostic AD, Clish CB, Vlamakis H, Xavier RJ, Huttenhower C. Predictive metabolomic profiling of microbial communities using amplicon or metagenomic sequences. Nat Commun. 2019;10(1):3136. doi: 10.1038/s41467-019-10927-1.
147. Le V, Quinn TP, Tran T, Venkatesh S. Deep in the bowel: highly interpretable neural encoder-decoder networks predict gut metabolites from gut microbiome. BMC Genomics. 2020;21(Suppl 4):256. doi: 10.1186/s12864-020-6652-7.
148. Reiman D, Layden BT, Dai Y. MiMeNet: exploring microbiome-metabolome relationships using neural networks. PLoS Comput Biol. 2021;17(5):e1009021. doi: 10.1371/journal.pcbi.1009021.
149. Morton JT, Aksenov AA, Nothias LF, Foulds JR, Quinn RA, Badri MH, Swenson TL, Van Goethem MW, Northen TR, Vazquez-Baeza Y, et al. Learning representations of microbe-metabolite interactions. Nat Methods. 2019;16(12):1306–1314. doi: 10.1038/s41592-019-0616-3.
150. Wang T, Wang X-W, Lee-Sarwar KA, Litonjua AA, Weiss ST, Sun Y, Maslov S, Liu Y-Y. Predicting metabolomic profiles from microbial composition through neural ordinary differential equations. Nat Mach Intell. 2023;5(3):284–293. doi: 10.1038/s42256-023-00627-3.
151. Weng Q, Hu M, Peng G, Zhu J. DMoVGPE: predicting gut microbial associated metabolites profiles with deep mixture of variational Gaussian process experts. BMC Bioinf. 2025;26(1):93. doi: 10.1186/s12859-025-06110-7.
152. Wolever TM, Jenkins DJ, Ocana AM, Rao VA, Collier GR. Second-meal effect: low-glycemic-index foods eaten at dinner improve subsequent breakfast glycemic response. AJCN. 1988;48(4):1041–1047. doi: 10.1093/ajcn/48.4.1041.
153. National Institutes of Health (NIH). (2020) 2020–2030 Strategic Plan for NIH Nutrition Research [Internet]. [place unknown]: U.S. Department of Health and Human Services. https://dpcpsi.nih.gov/sites/g/files/mnhszr346/files/2020NutritionStrategicPlan_508.pdf.
154. Dang T, Lysenko A, Boroevich KA, Tsunoda T. VBayesMM: variational Bayesian neural network to prioritize important relationships of high-dimensional microbiome multiomics data. Brief Bioinform. 2025;26(4):bbaf300. doi: 10.1093/bib/bbaf300.

155. Giuffrè M, Moretti R, Tiribelli C. Gut microbes meet machine learning: the next step towards advancing our understanding of the gut microbiome in health and disease. Int J Mol Sci. 2023;24(6):5229. doi: 10.3390/ijms24065229.
156. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. Nat Rev Genet. 2014;15(2):107–120. doi: 10.1038/nrg3643.
157. Mendoza SN, Olivier BG, Molenaar D, Teusink B. A systematic assessment of current genome-scale metabolic reconstruction tools. Genome Biol. 2019;20(1):158. doi: 10.1186/s13059-019-1769-1.
158. Machado D, Andrejev S, Tramontano M, Patil KR. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. Nucleic Acids Res [Internet]. 2018;46:7542–7553. doi: 10.1093/nar/gky537.
159. Kim M, Sung J, Chia N. Resource-allocation constraint governs structure and function of microbial communities in metabolic modeling. Metab Eng. 2022;70:12–22. doi: 10.1016/j.ymben.2021.12.011.
160. Heinken A, Basile A, Thiele I. Advances in constraint-based modelling of microbial communities. Curr Opin Syst Biol. 2021;27:100346100346. doi: 10.1016/j.coisb.2021.05.007.
161. Quinn-Bohmann N, Carr AV, Diener C, Gibbons SM. Moving from genome-scale to community-scale metabolic models for the human gut microbiome. Nat Microbiol. 2025;10(5):1055–1066. doi: 10.1038/s41564-025-01972-2.
162. Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, Greenhalgh K, Jäger C, Baginska J, Wilmes P, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat Biotechnol. 2017;35(1):81–89. doi: 10.1038/nbt.3703.
163. Heinken A, Hertel J, Acharya G, Ravcheev DA, Nyga M, Okpala OE, Hogan M, Magnúsdóttir S, Martinelli F, Nap B, et al. Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized Medicine. Nat Biotechnol. 2023;41(9):1320–1331. doi: 10.1038/s41587-022-01628-0.
164. Noronha A, Modamio J, Jarosz Y. The virtual metabolic human database: integrating human and gut microbiome metabolism with nutrition and disease. Nucleic Acids Res [Internet]. 2018;47:D614–D624. doi: 10.1093/nar/gky992.
165. Orth JD, Thiele I, Palsson B. What is flux balance analysis? Nat Biotechnol. 2010;28(3):245–248. doi: 10.1038/nbt.1614.
166. Gudmundsson S, Thiele I. Computationally efficient flux variability analysis. BMC Bioinf. 2010;11(1):489. doi: 10.1186/1471-2105-11-489.
167. Dukovski I, Bajić D, Chacón JM, Quintin M, Vila JCC, Sulheim S, Pacheco AR, Bernstein DB, Riehl WJ, Korolev KS, et al. A metabolic modeling platform for the computation of microbial ecosystems in time and space (COMETS). Nat Protoc. 2021;16(11):5030–5082. doi: 10.1038/s41596-021-00593-3.
168. Bauer E, Zimmermann J, Baldini F, Thiele I, Kaleta C. BacArena: individual-based metabolic modeling of heterogeneous microbes in complex communities. PLoS Comput Biol. 2017;13(5):e1005544. doi: 10.1371/journal.pcbi.1005544.
169. Zomorrodi AR, Islam MM, Maranas CD. d-OptCom: dynamic multi-level and multi-objective metabolic modeling of microbial communities. ACS Synth Biol. 2014;3(4):247–257. doi: 10.1021/sb4001307.
170. Biggs MB, Papin JA. Novel multiscale modeling tool applied to pseudomonas aeruginosa biofilm formation. PLoS One. 2013;8(10):e78011. doi: 10.1371/journal.pone.0078011.
171. Choudhary R, Mahadevan R. DyMMM-LEAPS: an ML-based framework for modulating evenness and stability in synthetic microbial communities. Biophys J. 2024;123(18):2974–2995. doi: 10.1016/j.bpj.2024.05.006.
172. Louca S, Doebeli M. Calibration and analysis of genome-based models for microbial ecology. eLife. 2015;4:e08208. doi: 10.7554/eLife.08208.
173. Shoaie S, Ghaffari P, Kovatcheva-Datchary P, Mardinoglu A, Sen P, Pujos-Guillot E, de Wouters T, Juste C, Rizkalla S, Chilloux J, et al. Quantifying diet-induced metabolic changes of the human gut microbiome. Cell Metab. 2015;22(2):320–331. doi: 10.1016/j.cmet.2015.07.001.
174. Quinn-Bohmann N, Wilmanski T, Sarmiento KR, Levy L, Lampe JW, Gurry T, Rappaport N, Ostrem EM, Venturelli OS, Diener C, et al. Microbial community-scale metabolic modelling predicts personalized short-chain fatty acid production profiles in the human gut. Nat Microbiol. 2024;9(7):1700–1712. doi: 10.1038/s41564-024-01728-4.
175. Diener C, Gibbons SM, Resendis-Antonio O. MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. mSystems. 2020;5(1):e00606-19. doi: 10.1128/mSystems.00606-19.
176. Zampieri G, Vijayakumar S, Yaneske E, Angione C. Machine and deep learning meet genome-scale metabolic modeling. PLoS Comput Biol. 2019;15(7):e1007084. doi: 10.1371/journal.pcbi.1007084.
177. Faure L, Mollet B, Liebermeister W, Faulon J-L. A neural-mechanistic hybrid approach improving the predictive power of genome-scale metabolic models. Nat Commun. 2023;14(1):4669. doi: 10.1038/s41467-023-40380-0.
178. van der Ark KCH, van Heck RGA, Martins Dos Santos VAP, Belzer C, de Vos WM. More than just a gut feeling: constraint-based genome-scale metabolic models for predicting functions of human intestinal microbes. Microbiome. 2017;5(1):78. doi: 10.1186/s40168-017-0299-x.

179. Babaei P, Shoaie S, Ji B, Nielsen J. Challenges in modeling the human gut microbiome. Nat Biotechnol. 2018;36(8):682–686. doi: 10.1038/nbt.4213.
180. Paulson JA, Martin-Casas M, Mesbah A. Fast uncertainty quantification for dynamic flux balance analysis using non-smooth polynomial chaos expansions. PLoS Comput Biol. 2019;15(8):e1007308. doi: 10.1371/journal.pcbi.1007308.
181. Sen P, Orešič M. Integrating omics data in genome-scale metabolic modeling: a methodological perspective for precision Medicine. Metabolites. 2023;13(7):855. doi: 10.3390/metabo13070855.
182. Papoutsoglou G, Tarazona S, Lopes MB, Klammsteiner T, Ibrahimi E, Eckenberger J, Novielli P, Tonda A, Simeon A, Shigdel R, et al. Machine learning approaches in microbiome research: challenges and best practices. Front Microbiol. 2023;14:1261889. doi: 10.3389/fmicb.2023.1261889.
183. Holmes ZC, Villa MM, Durand HK, Jiang S, Dallow EP, Petrone BL, Silverman JD, Lin P-H, David LA. Microbiota responses to different prebiotics are conserved within individuals and associated with habitual fiber intake. Microbiome. 2022;10(1):114. doi: 10.1186/s40168-022-01307-x.
184. Williams CE, Hammer TJ, Williams CL. Diversity alone does not reliably indicate the healthiness of an animal microbiome. ISME J. 2024;18(1):wrae133. doi: 10.1093/ismejo/wrae133.
185. Shanahan F, Ghosh TS, O'Toole PW. The healthy microbiome-what is the definition of a healthy gut microbiome? Gastroenterology. 2021;160(2):483–494. doi: 10.1053/j.gastro.2020.09.057.
186. Gupta VK, Kim M, Bakshi U, Cunningham KY, Davis JM, 3rd, Lazaridis KN, Nelson H, Chia N, Sung J. A predictive index for health status using species-level gut microbiome profiling. Nat Commun. 2020;11(1):4635. doi: 10.1038/s41467-020-18476-8.
187. Chang D, Gupta VK, Hur B, Cobo-López S, Cunningham KY, Han NS, Lee I, Kronzer VL, Teigen LM, Karnatovskaia LV, et al. Gut microbiome wellness index 2 enhances health status prediction from gut microbiome taxonomic profiles. Nat Commun. 2024;15(1):7447. doi: 10.1038/s41467-024-51651-9.
188. Zhu J, Xie H, Yang Z, Chen J, Yin J, Tian P, Wang H, Zhao J, Zhang H, Lu W, et al. Statistical modeling of gut microbiota for personalized health status monitoring. Microbiome. 2023;11(1):184. doi: 10.1186/s40168-023-01614-x.
189. Nguyen SM, Tran HTT, Long J, Shrubsole MJ, Cai H, Yang Y, Nguyen LM, Nguyen GH, Nguyen CV, Ta TV, et al. Gut microbiome of patients with breast cancer in Vietnam. JCO Glob Oncol. 2024;10(10):e2300234. doi: 10.1200/GO.23.00234.
190. Tap J, Lejzerowicz F, Cotillard A, Pichaud M, McDonald D, Song SJ, Knight R, Veiga P, Derrien M. Global branches and local states of the human gut microbiome define associations with environmental and intrinsic factors. Nat Commun. 2023;14(1):3310. doi: 10.1038/s41467-023-38558-7.
191. Liu X, Zhou H, Zhang J, Li R, Liang J. Brain-gut co-management: probiotic LAB improves mental health and further reduces disease activity in ulcerative colitis patients with emotional disturbance. Nutr Neurosci. 2025;28(12):1511–1522. doi: 10.1080/1028415X.2025.2527224.
192. Li Z, Ye Y, Lai S, Bao J, Fu A. Puerarin affects adipose lipolysis and ameliorates obesity by gut microbiota. Front Microbiol. 2025;16:1567339. doi: 10.3389/fmicb.2025.1567339.
193. Gacesa R, Kurilshikov A, Vich Vila A, Sinha T, Klaassen MAY, Bolte LA, Andreu-Sánchez S, Chen L, Collij V, Hu S, et al. Environmental factors shaping the gut microbiome in a Dutch population. Nature. 2022;604(7907):732–739. doi: 10.1038/s41586-022-04567-7.
194. Yin X, Tong Q, Wang J, Wei J, Qin Z, Wu Y, Zhang R, Guan B, Qiu H. The impact of altered dietary adenine concentrations on the gut microbiota in drosophila. Front Microbiol. 2024;15:1433155. doi: 10.3389/fmicb.2024.1433155.
195. Lee DH, Seong H, Chang D, Gupta VK, Kim J, Cheon S, Kim G, Sung J, Han NS. Evaluating the prebiotic effect of oligosaccharides on gut microbiome wellness using in vitro fecal fermentation. Npj Sci Food. 2023;7(1):18. doi: 10.1038/s41538-023-00195-1.
196. Gupta VK, Rajendraprasad S, Ozkan M, Ramachandran D, Ahmad S, Bakken JS, Laudanski K, Gajic O, Bauer B, Zec S, et al. Safety, feasibility, and impact on the gut microbiome of kefir administration in critically ill adults. BMC Med. 2024;22(1):80. doi: 10.1186/s12916-024-03299-x.
197. Zhang Z, Yang Z, Lin S, Jiang S, Zhou X, Li J, Lu W, Zhang J. Probiotic-induced enrichment of adlercreutzia equolifaciens increases gut microbiome wellness index and maps to lower host blood glucose levels. Gut Microbes. 2025;17(1):2520407. doi: 10.1080/19490976.2025.2520407.
198. Yang Z, Xie H, Zhu J, Wang H, Lu W. Development of a dietary factor evaluation method based on the gut microbiota health index. Sheng Wu Gong Cheng Xue Bao. 2025;41(6):2373–2387.
199. Gupta VK, Paul S, Dutta C. Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. Front Microbiol. 2017;8:1162. doi: 10.3389/fmicb.2017.01162.
200. Patangia DV, Anthony Ryan C, Dempsey E, Paul Ross R, Stanton C. Impact of antibiotics on the human microbiome and consequences for host health. Microbiologyopen. 2022;11(1):e1260. doi: 10.1002/mbo3.1260.
201. Nalbantoglu OU, Ermis BH, Gundogdu A. Deep learning-enhanced wellness scores: a population level study on gut microbiome profiling and health prediction. Biomed Signal Process Control. 2025;110:108146108146. doi: 10.1016/j.bspc.2025.108146.

202. Kase BE, Liese AD, Zhang J, Murphy EA, Zhao L, Steck SE. The development and evaluation of a literature-based dietary index for gut microbiota. Nutrients. 2024;16(7):1045. doi: 10.3390/nu16071045.
203. Wang C, Segal LN, Hu J, Zhou B, Hayes RB, Ahn J, Li H. Microbial risk score for capturing microbial characteristics, integrating multi-omics data, and predicting disease risk. Microbiome. 2022;10(1):121. doi: 10.1186/s40168-022-01310-2.
204. Blount K, Jones C, Walsh D, Gonzalez C, Shannon WD. Development and validation of a novel microbiome-based biomarker of post-antibiotic dysbiosis and subsequent restoration. Front Microbiol. 2021;12:781275. doi: 10.3389/fmicb.2021.781275.
205. Subramanian S, Huq S, Yatsunenko T, Haque R, Mahfuz M, Alam MA, Benezra A, DeStefano J, Meier MF, Muegge BD, et al. Persistent gut microbiota immaturity in malnourished Bangladeshi children. Nature. 2014;510(7505):417–421. doi: 10.1038/nature13421.
206. Sczyrba A, Hofmann P, Belmann P, Koslicki D, Janssen S, Dröge J, Gregor I, Majda S, Fiedler J, Dahms E, et al. Critical assessment of metagenome Interpretation-a benchmark of metagenomics software. Nat Methods. 2017;14(11):1063–1071. doi: 10.1038/nmeth.4458.
207. Meyer F, Robertson G, Deng Z-L, Koslicki D, Gurevich A, McHardy AC. CAMI benchmarking portal: online evaluation and ranking of metagenomic software. Nucleic Acids Res. 2025;53(W1):W102–W109. doi: 10.1093/nar/gkaf369.
208. Rappaport N, Nogal B, Perrott K, Domina V, Hood L, Price ND. Early detection of wellness-to-disease transitions in the AI era: implications for pharmacology and toxicology. Annu Rev Pharmacol Toxicol. 2026;66(1):41–64. doi: 10.1146/annurev-pharmtox-062124-013423.
209. Rappaport N, Schweickart A, Hood L, Price ND. We wait for disease to shout-what if we listened when biology whispered? Cell Syst. 2026;17(2):101509. doi: 10.1016/j.cels.2025.101509.

Appendix

Box 1. Glossary of key artificial intelligence and machine learning terms in precision nutrition and diet-gut microbiome research. This glossary summarizes key artificial intelligence (AI), machine learning (ML), and modeling concepts used throughout this review, with emphasis on their application to integrating dietary, microbial, and host data in precision nutrition.

1. **Learning Paradigms**
   - **Artificial Intelligence (AI)**: A broad field of computer science aimed at creating systems that perform tasks requiring human-like intelligence, such as pattern recognition, decision-making, and prediction. In diet-gut microbiome research, AI encompasses ML and other techniques to analyze complex interactions (e.g., predicting glycemic responses from gut microbial profiles).
     - **Example:** AI-guided personalization in randomized controlled trials, where algorithms recommend diets based on multi-omics data, yielding metabolic benefits like improved blood glucose control.
     - **Nuances and Implications:** AI is often conflated with ML (a subset), but includes rule-based systems. Ethical implications include bias in training data (e.g., underrepresentation of non-Western populations), potentially leading to inequitable nutrition advice.
     - **Considerations:** Scalability for clinical use requires interpretable AI to build trust among practitioners.
   - **Machine Learning (ML)**: A subset of AI in which algorithms learn patterns from data without explicit rule-based programming. ML enables prediction of metabolic responses, diet classification, and responder stratification from high-dimensional microbiome data.
     - **Example:** Using ML to classify habitual diets from shotgun metagenomic signatures or stratify responders to probiotic interventions.
     - **Nuances and Implications:** ML can be supervised (labeled data, e.g., healthy vs. diseased microbiomes) or unsupervised (no labels, e.g., clustering microbial communities). Overfitting (model memorizing noise) is a risk in sparse omics data; regularization techniques like LASSO mitigate this.
     - **Considerations:** Generalizability across cohorts is key; diverse longitudinal datasets (e.g., from multiple continents) reduce biases.
   - **Supervised learning**: Models trained on labeled input-output pairs (e.g., microbiome features linked to glycemic response). Tasks include classification and regression. Common algorithms include regularized regression, random forests, gradient boosting, support vector machines, and neural networks.
   - **Unsupervised learning**: Models that identify structure in unlabeled data, used for clustering (e.g., microbiome states), dimensionality reduction, and latent pattern discovery.
   - **Semi-supervised and weakly supervised learning**: Approaches that leverage limited or noisy labels (e.g., self-reported diet) alongside abundant unlabeled omics data, improving sample efficiency.
   - **Multi-task learning**: A framework in which a single model simultaneously predicts multiple related outcomes (e.g., glycemic, lipid, and inflammatory responses), sharing internal representations to enhance generalization.
2. **Model Architectures and Representation**

- **Ensemble methods**: Techniques that combine multiple base models to improve predictive accuracy and robustness. Examples include bagging (random forests), boosting (gradient-boosted decision trees), and stacking.
- **Deep Learning (DL)**: A subset of ML using multi-layer neural networks to learn hierarchical representations directly from high-dimensional inputs (e.g., metagenomes, metabolomes, continuous glucose traces). DL enables end-to-end multimodal integration but often requires large datasets.
  - **Example:** DL models in Figure 1 for digital twin simulations, predicting metabolic outcomes from metagenomic and dietary inputs.
  - **Nuances and Implications:** Requires large datasets and computational power; "black-box" nature reduces interpretability, but techniques like attention mechanisms help. Implications include improved accuracy (e.g., 20–30% better glycemic forecasting) but higher risk of overfitting in small cohorts.
  - **Considerations:** Hybrid DL-mechanistic models balance data-driven power with biological plausibility.
- **Representation learning and embeddings**: Methods that learn informative lower-dimensional feature spaces from complex inputs. Examples include autoencoders, word2vec-style embeddings for microbe-metabolite relationships, and transformer-based latent representations for microbiome health metrics.

3. **Data Handling and Preprocessing**
   - **Multi-omics profiling**: Integration of multiple molecular layers (metagenomics, metatranscriptomics, metaproteomics, metabolomics) to capture complementary aspects of microbial structure and function.
     - **Example:** Combining shotgun metagenomics with host clinical biomarkers to predict individual responses to dietary perturbations.
     - **Nuances and Implications:** Addresses data heterogeneity (e.g., varying scales across omics) but introduces high dimensionality (thousands of features). Enables finer resolution but risks information overload; preprocessing is crucial.
     - **Considerations:** Standardization (e.g., via shared protocols) is needed for cross-study comparability.
   - **High dimensionality and sparsity**: Microbiome datasets typically contain thousands of features relative to sample size, with many low-abundance or zero-inflated taxa.
   - **Compositional data**: Relative abundance data constrained by constant-sum properties; requires specialized transformations (e.g., centered log-ratio) to avoid spurious correlations.
   - **Feature engineering and feature selection**: Processes that transform raw inputs into model-ready variables and identify informative subsets. Methods include domain-driven aggregation, LASSO-based selection, and tree-based importance metrics.
     - **Example:** LASSO-penalized logistic regression in GMWI2 selects discriminative taxa for wellness scores.
     - **Nuances and Implications:** Combats the "curse of dimensionality" in multi-omics, improving model performance. Biological interpretability (e.g., highlighting key microbes like *Faecalibacterium prausnitzii*).
     - **Considerations:** Domain knowledge (e.g., microbial ecology) guides selection to avoid data-driven artifacts.
   - **Dimensionality reduction**: Techniques (e.g., PCA, t-SNE, UMAP, autoencoders) that compress high-dimensional data while retaining essential structure.
     - **Example:** In hiPCA (Health Index with PCA), reducing taxonomic features to identify overweight-associated microbiome states, showing decreases post-dietary interventions like galacto-oligosaccharides.
     - **Nuances and Implications:** Preserves variance but can lose rare signals (e.g., low-abundance taxa). Improves model efficiency but may introduce biases if assumptions (e.g., linearity in PCA) don't hold.
     - **Considerations:** Use in conjunction with imputation for sparse microbiome datasets.
   - **Data augmentation, imputation, normalization/scaling**: Preprocessing steps that address missing data, heterogeneous measurement scales, and class imbalance in multimodal datasets.
     - **Example:** In Figure 1 workflows, imputing missing metabolomic peaks or normalizing taxonomic abundances for ML integration.
     - **Nuances and Implications:** Essential for heterogeneous inputs (e.g., self-reported diets vs. NGS data). Reduces bias but can introduce errors (e.g., mean imputation assuming normality).
     - **Considerations:** Context-specific (e.g., compositionality-aware normalization for microbiome data).
4. **Model Evaluation and Generalization**
   - **Overfitting**: When a model captures dataset-specific noise rather than generalizable patterns, leading to poor external performance.
   - **Regularization**: Techniques (e.g., L1/LASSO, L2/Ridge, dropout, early stopping) that constrain model complexity to mitigate overfitting.
     - **Example:** In GMWI2 construction, using LASSO to select health-associated taxa from 8,069 metagenomes, outputting wellness scores.
     - **Nuances and Implications:** Balances bias-variance tradeoff; interpretable via selected features. Translates high-dimensional data into actionable metrics outperforming α-diversity.
     - **Considerations:** Parameter tuning (e.g., via cross-validation) is critical for microbiome applications.
   - **Cross-validation and nested cross-validation**: Resampling frameworks used to assess generalization and tune hyperparameters without optimistic bias.

- **External validation**: Evaluation on independent cohorts to assess robustness across populations, sequencing platforms, and dietary contexts.
- **Domain shift**: Performance degradation when applying a model to data from different demographic, geographic, or technical contexts than those used in training.
- **Uncertainty estimation**: Approaches (e.g., Bayesian inference, ensemble variance, conformal prediction) that quantify prediction confidence. Critical for clinical deployment.

5. **Interpretability and Causality**
   - **Interpretability and explainability**: Methods that clarify why a model produces a given prediction. Examples include permutation importance, SHAP values, LIME, saliency maps, and attention mechanisms.
   - **Black-box models**: High-capacity models (e.g., deep neural networks) whose internal decision processes are not directly transparent.
   - **Causal inference vs. associative prediction**: Most ML models identify statistical associations rather than causal effects. Integrating causal frameworks or mechanistic constraints helps distinguish predictive correlates from modifiable drivers.
6. **Mechanistic and Hybrid Modeling**
   - **Genome-Scale Metabolic Models (GSMMs)**: Constraint-based mechanistic models that simulate metabolic fluxes within microbial communities or hosts, based on genome annotations.
     - **Example:** *In silico* simulations of dietary perturbations, guiding targeted interventions like prebiotics.
     - **Nuances and Implications:** Hypothesis-driven, contrasting data-driven ML; provides biological plausibility. Enhances interpretability but requires accurate annotations.
     - **Considerations:** Integration with ML yields hybrid models for scalable predictions.
   - **Flux Balance Analysis (FBA)**: Optimization-based method used in GSMMs to estimate feasible steady-state metabolic flux distributions.
   - **Digital twins**: Individualized *in silico* models of a person's microbiome or host-microbe system used to simulate responses to dietary or probiotic interventions.
   - **Hybrid AI-mechanistic models**: Frameworks combining ML prediction with mechanistic constraints (e.g., GSMM-derived flux features or differentiable FBA surrogates) to enhance biological plausibility, interpretability, and generalizability.
     - **Example:** Merging ML for pattern detection with GSMMs for metabolic insights, improving generalizability in precision nutrition.
     - **Nuances and Implications:** Addresses ML's lack of causality and mechanisms' data limitations. Enhances scalability and interpretability (e.g., explaining why a diet boosts GMWI2).
     - **Considerations:** Promising for translational challenges, like closed-loop optimization.
7. **Translational Concepts**
   - **Responder stratification**: ML-based classification of individuals into likely responders or non-responders to dietary interventions.
     - **Example:** Using microbiome signatures to identify individuals benefiting from AI-guided diets, as in trials showing metabolic gains.
     - **Nuances and Implications:** Enables personalized interventions; relies on robust features. Reduces trial failures but risks over-stratification.
     - **Considerations:** Validates across diverse cohorts for equity.
   - ***In silico* simulation**: Computational modeling of dietary perturbations or probiotic effects without real-world experimentation.
     - **Example:** Simulating probiotic effects on community metabolism to predict GMWI2 shifts. Digital Twins (virtual replicas for personalized simulations).
     - **Nuances and Implications:** Cost-effective for hypothesis testing; limited by model assumptions. Guides trials but needs empirical validation.
     - **Considerations:** Integrates with ML for adaptive recommendations.
   - **Composite microbiome health metrics**: Machine learning-derived indices (e.g., GMWI2, hiPCA, DI-GM) that summarize high-dimensional microbial data into interpretable wellness scores.
     - **Example:** GMWI2 and DI-GM for monitoring diet-induced wellness, correlating with lower glucose in interventions.
     - **Nuances and Implications:** Actionable for clinicians; context-dependent (e.g., diet-specific). Outperforms traditional metrics but requires large training data.
     - **Considerations:** Longitudinal tracking enhances utility in epidemiology.
   - **Closed-loop optimization systems**: Adaptive frameworks that iteratively refine dietary recommendations based on real-time data streams (e.g., continuous glucose monitoring, periodic microbiome sampling).
   - **Software as a Medical Device (SaMD)**: Regulatory classification for AI-driven clinical decision-support systems intended for medical use.