Explainable Artificial Intelligence-Guided Discovery of Plant Polysaccharide–Gut Microbiota Interactions for Precision Intervention in Metabolic Syndrome

Zhoutianyi Liu
Department of Computer Science and Engineering, University of Nevada, Reno, Reno, NV, USA.
zliu@unr.edu

Anel Arera
Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS, USA.
arora934@ku.edu

Isaac R. Carpenter
School of Computing, Clemson University, Clemson, SC, USA.
carpenterisaac@clemson.edu

Zeixan Xia
Department of Computer Science, University of North Texas, Denton, TX, USA.
hellozixuan@unt.edu

Abstract

Metabolic syndrome represents a global health crisis in which gut microbiota dysbiosis plays a central role, and plant-derived polysaccharides have emerged as promising modulators capable of restoring microbial equilibrium and improving host metabolic health. However, the immense structural diversity of plant polysaccharides, the inter-individual variability of gut microbial communities, and the nonlinear, multi-layered nature of host–microbe interactions render conventional discovery and intervention paradigms insufficient for precision applications. This paper presents a systems-level framework that leverages explainable artificial intelligence to guide the discovery of plant polysaccharide–gut microbiota interactions, with the explicit goal of enabling precision nutritional interventions against metabolic syndrome. We examine the architectural requirements for integrating multi-omics data, polysaccharide structural descriptors, and clinical phenotypes within a transparent machine learning pipeline. The discussion spans data infrastructure, model interpretability methods such as feature attribution and concept-based explanations, deployment architectures for clinical decision support, and the governance and policy dimensions that must accompany the translation of algorithmic insights into real-world dietary strategies. By situating explainable artificial intelligence as an orchestrating layer across molecular, microbial, and population scales, the paper articulates how structural trade-offs in model complexity, data privacy, fairness, sustainability, and regulatory accountability shape the design of trustworthy systems. The analysis reveals that explainability is not merely a technical add-on but a foundational prerequisite for building credible causal models, enabling reproducible mechanistic inference, and earning acceptance from clinicians, regulators, and diverse populations. We conclude by identifying cross-domain synergies and future directions that

integrate ecological modeling, federated learning, and policy frameworks to create a robust pipeline from polysaccharide characterization to culturally adapted, evidence-based metabolic syndrome management.

Keywords

explainable artificial intelligence, gut microbiota, plant polysaccharides, metabolic syndrome, precision intervention, multi-omics integration, interpretable machine learning, systems architecture.

1. Introduction

Metabolic syndrome constitutes a cluster of interconnected cardiometabolic abnormalities—central obesity, dyslipidemia, hypertension, and impaired glucose tolerance—that collectively elevate the risk of type 2 diabetes, cardiovascular disease, and premature mortality [1]. The global prevalence has risen relentlessly, driven by dietary shifts toward energy-dense but fiber-depleted foods and increasingly sedentary lifestyles. Contemporary therapeutic strategies, although partially effective, remain insufficient to reverse the pandemic trajectory, prompting a search for interventions that target upstream etiological pathways. Among these, the gut microbiota has emerged as a critical nexus linking dietary inputs to host metabolic regulation. Seminal work established that an altered microbial ecosystem, characterized by reduced community richness and shifts in the ratio of potentially beneficial to pro-inflammatory taxa, directly contributes to metabolic endotoxemia, altered energy harvest, and chronic low-grade inflammation [2]. This recognition has motivated attempts to intentionally remodel the gut microbiome through dietary components, particularly plant-derived polysaccharides that serve as substrates for microbial fermentation.

Plant polysaccharides encompass an extraordinary array of structurally heterogeneous macromolecules, including cellulose, hemicelluloses, pectins, fructans, resistant starches, and numerous species-specific heteropolysaccharides. Their degradation by specialized gut bacterial consortia yields short-chain fatty acids and other metabolites that modulate host appetite, insulin sensitivity, and lipid metabolism through G-protein-coupled receptors and epigenetic mechanisms [3]. The conceptual appeal of harnessing these natural fibers as targeted prebiotic interventions is immense, yet translation to precision medicine remains profoundly challenging. Individual responses to the same polysaccharide are highly variable, dependent on baseline microbiome composition, host genetics, and environmental factors. Conventional reductionist approaches that isolate a single structural feature and test it in a uniform population cannot capture the combinatorial complexity that governs ecological and physiological outcomes. A paradigm shift is required, one that treats the polysaccharide–microbiota–host axis as a complex adaptive system whose behavior emerges from high-dimensional interactions across multiple scales.

Artificial intelligence, particularly deep learning, has demonstrated remarkable capacity to model high-dimensional biological data. However, early applications in microbiome science often relied on black-box models that achieved predictive accuracy at the expense of mechanistic opacity. In the context of nutritional interventions for metabolic syndrome, a prediction that a certain fiber blend will improve glycemic control in a subpopulation, without an explanation of which microbial taxa were enriched, which metabolic pathways were activated, and which polysaccharide motifs were responsible, provides limited actionable insight for refining formulations or guiding clinical decisions. Explainable artificial intelligence (XAI) addresses this gap by making model reasoning transparent to human

stakeholders. XAI methods can reveal which input features drive predictions, uncover nonlinear interactions among polysaccharide properties and microbial genes, and generate hypotheses about causal mechanisms that can be tested in subsequent experimental or clinical studies. The integration of XAI into the discovery pipeline thus transforms machine learning from a pattern-matching tool into a scientific instrument for causal inference and rational intervention design.

This paper advances a system-level perspective on XAI-guided discovery of plant polysaccharide–gut microbiota interactions for precision metabolic syndrome intervention. We examine the entire socio-technical architecture, from data acquisition and curation to model training, explanation generation, validation, and deployment within clinical and public health ecosystems. Rather than focusing narrowly on algorithmic performance, we emphasize structural trade-offs among model fidelity, interpretability, fairness, data privacy, and sustainability. The analysis spans data infrastructure for multi-omics integration, the selection and design of XAI methods suited to dynamic microbial networks, and the governance frameworks that must ensure equitable access and prevent algorithmic harm. By framing XAI as an orchestrating capability that weaves together molecular, microbial, and population-level information, we provide a blueprint for building trustworthy systems that can accelerate the discovery of polysaccharide-based prebiotics tailored to individual and population needs.

2. Metabolic Syndrome and the Gut Microbiota–Polysaccharide Axis

The pathophysiology of metabolic syndrome is deeply entwined with the gut microbial ecosystem. A diet rich in saturated fats and simple sugars, coupled with low intake of fermentable fibers, reduces the abundance of bacterial taxa that specialize in complex carbohydrate degradation, thereby diminishing the production of the short-chain fatty acids acetate, propionate, and butyrate. These metabolites are not merely passive end-products; butyrate serves as the primary energy source for colonocytes and strengthens intestinal barrier integrity, while propionate and acetate engage host receptors that regulate gluconeogenesis, lipogenesis, and appetite [1,2]. A compromised barrier allows lipopolysaccharides from Gram-negative bacteria to translocate into systemic circulation, triggering a cascade of inflammatory signals that impair insulin signaling in peripheral tissues. This mechanistic chain underscores the rationale for dietary interventions that replete fiber-degrading keystone species and restore a health-associated metabolic milieu.

Plant polysaccharides arrive in the colon as intact polymers that cannot be digested by host enzymes, thereby creating a competitive metabolic niche for bacteria endowed with the requisite carbohydrate-active enzymes. The structural complexity of these substrates—encompassing molecular weight, degree of branching, glycosidic linkage types, substitution patterns, and cross-linking with polyphenols—determines which microbial taxa can access them and which fermentation products are generated [3,4]. For example, long-chain inulin-type fructans preferentially stimulate Bifidobacterium species, whereas resistant starch type 4 may favor Ruminococcus bromii and related butyrate producers. The concept of prebiotics, formalized by an international consensus, stipulates that a substrate must be selectively utilized by host microorganisms to confer a health benefit [4]. Yet the selectivity is rarely absolute, and the health benefit is conditional on baseline ecological context. An intervention that enriches Bifidobacterium in one individual may enrich a different consortium in another due to cross-feeding networks and competitive exclusion dynamics that are not captured by simple dose-response curves.

The need for precision therefore arises from two interrelated sources of variability: the chemical variability of polysaccharides and the ecological variability of microbiota. Even a single plant species yields polysaccharides whose fine structure can change with harvest season, extraction method, and post-processing. Phyllostachys nigra, for instance, has been shown to yield a polysaccharide with notable glycolipid metabolism regulatory activity in murine models, an effect associated with restructuring of the gut microbiome [18]. Extending such findings from animal models to human populations requires a systematic mapping of how specific structural motifs interact with the genetic potential of an individual's microbial community. Without such mapping, the development of polysaccharide-based products risks becoming an empirical, trial-and-error endeavor with limited reproducibility across demographic groups. A systems approach must therefore integrate polysaccharide chemical fingerprints, metagenomic snapshots, and host metabolic readouts into a unified analytical framework.

3. The Promise and Complexity of Plant Polysaccharides as Microbiota-Directed Interventions

The chemical space of plant polysaccharides is vast, and this very vastness constitutes both their therapeutic promise and the principal obstacle to systematic discovery. Unlike small-molecule drugs with well-defined pharmacophores, polysaccharides exert their biological effects through multivalent interactions, physical gelation properties, and microbial metabolism that generates a cascade of secondary metabolites. The prebiotic index of a given polysaccharide is not an intrinsic property but an emergent outcome of its encounter with a particular microbial ecosystem. Consequently, structure–activity relationships that are predictable in homogeneous laboratory models often break down when confronted with real-world heterogeneity. Resistant starch from maize, for instance, shows divergent effects on glucose homeostasis depending on whether participants harbor a Prevotella-dominant or Bacteroides-dominant enterotype [5]. This ecological contingency implies that any discovery framework must accommodate context-dependent nonlinearities and be capable of learning from population data while retaining the capacity to personalize recommendations.

Meta-analyses of dietary fiber interventions confirm that fiber-rich diets improve multiple metabolic syndrome components, yet the effect sizes are modest and exhibit substantial inter-study heterogeneity [5]. Part of the heterogeneity arises because studies typically administer a single fiber type to heterogeneous cohorts without stratifying participants by baseline microbiome composition or polysaccharide degradation potential. The situation is analogous to pharmacogenomics, where genetic variants in drug-metabolizing enzymes determine efficacy and toxicity. Here, the functional capacity of an individual's microbiome—the collective set of carbohydrate-active enzyme genes and metabolic pathways—determines the metabolic fate of a polysaccharide. Recent advances in metagenomic sequencing and gene-centric analysis have made it possible to quantify this capacity, but the interpretive challenge remains formidable because thousands of microbial genes and dozens of host biomarkers must be jointly considered.

Furthermore, plant polysaccharides do not operate in isolation; they are consumed as components of whole foods or formulated into supplements that may include other bioactive compounds such as polyphenols, which themselves modulate the microbiome. This matrix effect complicates the attribution of health outcomes to specific polysaccharide features. A reductionist strategy that isolates a single polysaccharide for testing in a gnotobiotic mouse model can establish causal proof-of-concept but cannot fully replicate the complexity of

human dietary patterns. An XAI-driven system that ingests multi-factorial data from human intervention trials and observational cohorts can disentangle these entangled effects by learning to assign credit across a feature space that includes multiple polysaccharide descriptors, dietary covariates, and microbial functional profiles. The output is not simply a ranking of important features but a network of conditional dependencies that can be interrogated to generate mechanistic hypotheses suitable for targeted experimental validation.

4. Explainable Artificial Intelligence as a System-Level Discovery Framework

The application of artificial intelligence in drug discovery and healthcare delivery has accelerated rapidly, with algorithms now capable of predicting molecular properties, optimizing clinical trial designs, and personalizing treatment regimens [6]. However, the complexity of the polysaccharide–microbiota–host axis demands more than predictive accuracy; it demands transparency regarding the biological logic that underlies predictions. Explainable artificial intelligence encompasses a family of techniques that render the internal reasoning of machine learning models comprehensible to human experts. In the context of metabolic syndrome interventions, XAI can serve as a discovery engine that generates, refines, and prioritizes hypotheses about how specific polysaccharide structures interact with particular microbial taxa to influence host metabolic outcomes.

At the core of the XAI toolkit are post-hoc feature attribution methods that assign an importance score to each input variable for a given prediction. The SHAP (SHapley Additive exPlanations) framework, grounded in cooperative game theory, provides a unified measure of feature importance with desirable consistency properties [9]. When applied to a model that predicts postprandial glucose excursion from inputs that include polysaccharide dosage, molecular weight, branching density, and relative abundances of microbial genera, SHAP values can reveal, for a specific individual, that a high branching density is the primary driver of a predicted improvement because it cross-feeds a consortium of butyrate producers that are already abundant in that person’s gut. Such individual-level explanations are immediately clinically actionable and can be aggregated across a population to discover general principles. LIME (Local Interpretable Model-agnostic Explanations) offers a complementary approach by locally approximating a complex model with a simpler, interpretable surrogate, thereby highlighting the decisional boundaries in the feature space [10]. Both methods enable what-if reasoning: a clinician could ask how the prediction would change if the same individual were to ingest a pectin with lower methoxylation degree, stimulating interactive exploration of potential intervention strategies.

Beyond post-hoc explanations, inherently interpretable architectures such as attention-based models and concept bottleneck models have been proposed for biomedical applications. Graph neural networks, for example, can represent polysaccharide molecules as graphs of monosaccharide nodes connected by glycosidic bonds, encoding structural motifs that are recognized by specific microbial enzymes [11]. By inspecting attention weights, researchers can identify which substructures are most strongly associated with the enrichment of Akkermansia muciniphila or Faecalibacterium prausnitzii, bridging the gap between cheminformatics and microbial ecology. These interpretable representations can then be linked, through multi-task learning frameworks, to host metabolic outcomes such as changes in homeostatic model assessment of insulin resistance or serum triglyceride levels. The explanatory output thus traces a causal chain from molecular topology to microbial metabolism to host physiology, providing the granularity needed for rational prebiotic design.

A systems-level perspective emphasizes that XAI is not a post-hoc adornment but a design principle that must be embedded throughout the machine learning lifecycle. Data selection and preprocessing influence what patterns can be learned and later explained. For instance, if the taxonomic resolution is limited to genus level, XAI explanations cannot pinpoint species-level interactions that may be critical for metabolite production. If polysaccharide features are reduced to a single solubility parameter, the model cannot learn the role of specific linkage types. Consequently, the design of the feature space must be co-optimized with the choice of explanation method, involving interdisciplinary collaboration among carbohydrate chemists, microbiologists, and machine learning engineers. The entire pipeline becomes a socio-technical system in which domain knowledge shapes model construction and explanations, in turn, refine domain knowledge through iterative feedback loops.

5. Data Infrastructure and Multi-Omics Integration for Interaction Mapping

Building an XAI-guided discovery system for polysaccharide–microbiota interactions necessitates a robust data infrastructure capable of ingesting, harmonizing, and linking heterogeneous data types. The minimum required layers include high-resolution polysaccharide structural characterization data—monosaccharide composition, linkage analysis, molecular weight distribution, and spectroscopic fingerprints—as well as metagenomic, metatranscriptomic, and metabolomic profiles from stool samples. Host clinical data encompassing anthropometric measures, glycemic and lipid profiles, and inflammatory markers completes the picture. Population-scale studies such as those that have documented gut microbial associations with metabolic traits demonstrate the feasibility of generating these multi-omic layers at scale [7]. However, the integration of glycomic data with microbiome data remains a frontier challenge because standardized ontologies and data exchange formats for polysaccharide structures are less mature than those for genomes and proteomes.

Multi-omics integration methods have traditionally been classified into early, intermediate, and late fusion strategies. Early fusion concatenates all features into a single matrix, which is straightforward but can drown out subtle interactions in high-dimensional noise. Intermediate fusion uses separate encoders for each data modality, learning joint representations that capture cross-modal correlations, an approach that has been successfully applied to integrate host transcriptomes with gut metagenomes in inflammatory bowel disease [8]. For the polysaccharide–microbiota problem, an analogous architecture could encode polysaccharide molecular graphs through graph neural networks, encode microbial pathway abundance profiles through autoencoders, and then fuse the latent representations in a shared embedding space where metabolic outcomes are predicted. XAI techniques can then be applied not only to the final prediction layer but also to the latent representations, revealing whether the model has learned biologically meaningful axes of variation such as a gradient from high-fiber to low-fiber fermentative capacity. This architecture demands substantial computational resources and careful attention to batch effects, missing data, and temporal sampling inconsistencies. Federated learning frameworks that train models across multiple clinical sites without centralizing raw data can address privacy concerns while increasing the diversity and statistical power of the training set [16].

Beyond algorithmic integration, the governance of data ecosystems is paramount. Multi-omics data derived from human participants are sensitive, and the ability to re-identify individuals from microbiome sequences has been demonstrated, raising ethical and legal considerations that intersect with regulations such as the General Data Protection Regulation. Transparent consent processes, community engagement, and institutional review board

oversight are necessary, but not sufficient, to build public trust. The data infrastructure should incorporate technical mechanisms for data provenance tracking, versioning of preprocessed datasets, and auditability of model training runs to enable reproducibility and accountability. A decentralized data architecture where contributors retain custodianship of their data while contributing to model training through secure aggregation protocols embodies a governance model that aligns with principles of data sovereignty and equity. These structural choices carry downstream consequences for the quality and fairness of XAI explanations, because models trained on unrepresentative, biased datasets will produce explanations that may mislead clinical decision-making and exacerbate health disparities.

6. XAI-Driven Feature Attribution and Mechanistic Inference

The core analytical contribution of XAI in this domain lies in its capacity to bridge correlative machine learning and causal mechanistic reasoning. When a predictive model has been trained to estimate the six-month change in fasting glucose following a polysaccharide intervention, the application of SHAP or integrated gradients yields a vector of contributions for thousands of molecular and microbial features. Rather than treating these attributions as a final answer, a systems approach treats them as a high-dimensional map that guides deeper biological inquiry. A consistent positive attribution for the relative abundance of a butyrate-producing species, conditional on the presence of a specific arabinoxylan side-chain motif, may indicate a syntrophic relationship that can be further examined using metabolic modeling or in vitro co-culture experiments. The XAI output thus operationalizes the concept of abduction—inference to the best explanation—within a computational framework that is auditable and reproducible.

Case illustrations from related fields demonstrate the power of this approach. In non-alcoholic fatty liver disease, machine learning models trained on gut metagenomic features have identified microbial signatures that distinguish advanced fibrosis from mild disease, and subsequent inspection of feature importance pointed to specific metabolic pathways that were then validated in independent cohorts [12]. Extending this paradigm to polysaccharide interventions requires incorporating fiber-specific features into the model. One could envision a study in which participants with metabolic syndrome consume a standardized portfolio of structurally defined polysaccharides, and their pre- and post-intervention multi-omics profiles are used to train a gradient-boosted tree ensemble. The SHAP dependence plots would then display how the glycemic response changes as a function of, say, the relative abundance of Eubacterium rectale across different levels of pectin galacturonan backbone length. Such visualizations empower domain experts to interact with the model, to ask follow-up questions, and to design validation experiments that close the cycle between computation and bench science.

Crucially, XAI can also expose model failures and biases. If the same polysaccharide feature is attributed a strongly positive impact for one demographic group but a neutral or negative impact for another, this discrepancy may reflect either true biological heterogeneity—for instance, driven by differences in baseline microbiota composition shaped by long-term dietary patterns—or an artifact of unbalanced training data. Transparent reporting of subgroup explanations, analogous to disaggregated performance metrics required in algorithmic fairness audits, becomes a tool for scientific discovery rather than a compliance checkbox. It prompts investigations into whether a particular microbial enzyme gene cluster is less prevalent in certain ancestral populations, leading to differential utilization of a polysaccharide. Such insights are indispensable for designing inclusive prebiotic products and for avoiding the trap

of developing interventions that are optimized for a narrow segment of the global population. The interpretability of the model thus aligns the goals of scientific rigor and health equity [14].

7. System Architecture and Deployment for Precision Intervention

Translating an XAI-guided discovery pipeline into a deployable system for precision metabolic syndrome intervention requires careful architectural design that spans from laboratory characterization to point-of-care recommendation engines. At the back end, the system must interface with polysaccharide libraries that catalog structural features, purity, and functional assay data; with secure cloud storage for patient-derived multi-omics data; and with high-performance computing resources for model training and explanation generation. The middle layer comprises a modular machine learning platform that supports model versioning, automated explanation reports, and continuous integration of new data from completed clinical trials or real-world evidence. The front end is a clinician- and patient-facing interface that presents personalized polysaccharide recommendations along with interpretable rationales—such as "This arabinogalactan blend is expected to improve your insulin sensitivity because it promotes the growth of Roseburia species that are currently underrepresented in your gut, as indicated by your metagenomic profile and supported by mechanistic models."

Such a system raises profound deployment challenges that extend beyond software engineering. The integration of XAI into clinical workflows demands that explanations be calibrated to the cognitive load and decision-making needs of end-users. A primary care physician may require a concise summary linking a prebiotic recommendation to a patient's specific metabolic risk factors, whereas a dietitian may benefit from a more granular breakdown of the microbial metabolic pathways involved. Designing explanation interfaces that are both faithful to the underlying model and accessible to diverse stakeholders is an active research area at the intersection of human-computer interaction and biomedical informatics. Additionally, the system must accommodate temporal dynamics: the gut microbiome is not static but responds to dietary changes, medications, and diurnal rhythms. Longitudinal monitoring, enabled by at-home stool sampling kits and wearable glucose sensors, can feed into recurrent model updating, where each new data point refines the personalized recommendation. The architecture must therefore support streaming data, incremental model retraining, and real-time explanation refresh, all while maintaining strict data security protocols.

Deployment also engages regulatory pathways that are only beginning to adapt to AI-driven nutritional interventions. In many jurisdictions, prebiotic formulations are regulated as dietary supplements or functional foods, categories that do not require the rigorous efficacy and safety evidence expected of pharmaceuticals. However, when an AI system delivers a personalized medical recommendation for managing a diagnosed condition like metabolic syndrome, the line between general wellness advice and medical device software becomes blurred. The U.S. Food and Drug Administration's evolving framework for Software as a Medical Device and the European Union's Medical Device Regulation establish risk-based classifications that may apply to such systems if they perform automated analysis of patient data to generate therapeutic suggestions [20]. An XAI capability can be a strategic asset in this regulatory landscape, because it allows developers to demonstrate that recommendations are grounded in biologically plausible, auditable reasoning rather than opaque proprietary algorithms. Transparency documentation, including model cards and datasheets for datasets,

becomes part of the submission package, enabling regulators to assess safety, effectiveness, and potential biases.

8. Governance, Fairness, and Policy Implications

The governance of XAI-guided polysaccharide discovery systems must reconcile multiple competing values: scientific innovation, individual autonomy, population health equity, and the sustainable use of plant biodiversity. Algorithmic fairness concerns are acute because both the training data and the targeted health outcomes are shaped by social determinants. If the majority of multi-omics data are collected from affluent populations with access to high-fiber diets and healthcare, models may fail to recognize the microbial signatures of individuals enduring food insecurity, who may harbor distinct microbial adaptations. The use of biased models in clinical decision support could perpetuate a cycle in which advantaged groups receive more effective prebiotic recommendations while underserved groups are offered generic advice that yields inferior outcomes. Proactive fairness strategies include stratified data collection initiatives, the application of fairness-aware learning algorithms that penalize performance disparities across demographic subgroups, and post-hoc explanation audits that quantify differential feature attributions [14].

Data governance frameworks must also address the ownership and sharing of polysaccharide structural data derived from traditional plant knowledge. Many plant species with promising prebiotic polysaccharides are indigenous to specific regions and have been used in traditional medicine for centuries. The extraction and commercialization of these polysaccharides without benefit-sharing arrangements raise issues of biopiracy that have been partially addressed by the Nagoya Protocol but remain unresolved in the context of AI-driven discovery. A system that screens thousands of plant polysaccharides in silico and identifies a high-value candidate originating from a biodiversity hotspot must incorporate mechanisms to ensure that the benefits—whether financial, technological, or health-related—flow back to the communities and nations that stewarded the genetic resources. Policy frameworks can mandate the inclusion of provenance metadata in glycomic databases and require algorithmic impact assessments that consider these equity dimensions alongside technical performance metrics.

Sustainability considerations extend to agricultural practices and the environmental footprint of scaling up production of selected polysaccharides. The EAT–Lancet Commission has underscored that dietary shifts toward plant-based foods must operate within planetary boundaries [17]. If XAI identifies a specific polysaccharide from a slow-growing perennial plant as a highly effective prebiotic, rapid commercialization could lead to monocropping, habitat loss, and water resource strain unless agricultural policies and certification schemes proactively guide sustainable sourcing. An integrated governance approach would embed lifecycle assessment data into the polysaccharide library, allowing the recommendation engine to weigh environmental costs alongside health benefits. This multidimensional optimization aligns with the broader vision of One Health, recognizing that human metabolic health, microbial ecology, and ecosystem integrity are interdependent. Policy instruments such as green public procurement criteria for dietary supplements and sustainability labeling can incentivize manufacturers to adopt polysaccharides that score favorably on both efficacy and environmental metrics.

9. Conclusion

The convergence of plant polysaccharide chemistry, gut microbial ecology, and explainable artificial intelligence opens a transformative pathway toward precision interventions for metabolic syndrome. This paper has argued that realizing this potential requires a systems-level architecture in which XAI serves not as a peripheral interpretive tool but as the central organizing framework that connects multi-omics data, polysaccharide structural features, and clinical outcomes. The structural trade-offs inherent in this design—between model fidelity and interpretability, data centralization and privacy, individual personalization and population generalizability, commercial efficiency and equity—must be navigated through deliberate technical and governance choices. By rendering the inferential logic transparent, XAI enables a virtuous cycle of hypothesis generation, experimental validation, and model refinement that accelerates discovery while building the trust of clinicians, regulators, and diverse communities.

Looking ahead, several frontiers demand attention. The development of foundation models for microbiome and glycomic data, analogous to those that have revolutionized natural language processing, could provide rich contextualized representations that downstream XAI modules can interrogate. Coupling these representations with causal inference frameworks, such as instrumental variable analysis within observational datasets, may move beyond correlational attributions toward genuine causal effect estimation. On the deployment side, embedding XAI-based prebiotic recommendation engines within national digital health platforms could democratize access to personalized nutritional advice, provided that initiatives are accompanied by investments in underrepresented population data and inclusive design practices. Finally, the governance architecture must evolve in step with technological capabilities, establishing international norms for data sharing, algorithmic transparency, and equitable benefit distribution that reflect the globally interconnected nature of both the gut microbiome and the plant biodiversity upon which it depends. The interdisciplinary collaboration required is vast, but the societal burden of metabolic syndrome and the promise of harnessing nature's molecular diversity demand that we undertake this effort with rigor, urgency, and ethical commitment.

References

1. Eckel, R. H., Grundy, S. M., & Zimmet, P. Z. (2005). The metabolic syndrome. The Lancet, 365(9468), 1415–1428.

2. Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., & Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature, 444(7122), 1027–1031.

3. Cockburn, D. W., & Koropatkin, N. M. (2016). Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. Journal of Molecular Biology, 428(16), 3230–3252.

4. Gibson, G. R., Hutkins, R., Sanders, M. E., Prescott, S. L., Reimer, R. A., Salminen, S. J., Scott, K., Stanton, C., Swanson, K. S., Cani, P. D., Verbeke, K., & Reid, G. (2017). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. Nature Reviews Gastroenterology & Hepatology, 14(8), 491–502.

5. Reynolds, A. N., Akerman, A. P., & Mann, J. (2020). Dietary fibre and whole grains in diabetes management: Systematic review and meta-analyses. PLoS Medicine, 17(3), e1003053.

6. Liu, S. (2025). Revolutionizing healthcare: The role of artificial intelligence in drug discovery and delivery. Journal of Angiotherapy, 9(1), 1-16.

7. Zhernakova, A., Kurilshikov, A., Bonder, M. J., Tigchelaar, E. F., Schirmer, M., Vatanen, T., Mujagic, Z., Vila, A. V., Falony, G., Vieira-Silva, S., Wang, J., Imhann, F., Brandsma, E., Jankipersadsing, S. A., Joossens, M., Cenit, M. C., Deelen, P., Swertz, M. A., Weersma, R. K., … Fu, J. (2016). Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. Science, 352(6285), 565–569.

8. Lloyd-Price, J., Arze, C., Ananthakrishnan, A. N., Schirmer, M., Avila-Pacheco, J., Poon, T. W., Andrews, E., Ajami, N. J., Bonham, K. S., Brislawn, C. J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T. G., Hall, A. B., Lake, K., Landers, C. J., Mallick, H., Plichta, D. R., … Huttenhower, C. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature, 569(7758), 655–662.

9. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

10. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144.

11. Stokes, J. M., Yang, K., Swanson, K., Jin, W., Cubillos-Ruiz, A., Donghia, N. M., MacNair, C. R., French, S., Carfrae, L. A., Bloom-Ackermann, Z., Tran, V. M., Chiappino-Pepe, A., Badran, A. H., Andrews, I. W., Chory, E. J., Church, G. M., Brown, E. D., Jaakkola, T. S., Barzilay, R., & Collins, J. J. (2020). A deep learning approach to antibiotic discovery. Cell, 180(4), 688–702.

12. Loomba, R., Seguritan, V., Li, W., Long, T., Klitgord, N., Bhatt, A., Dulai, P. S., Caussy, C., Bettencourt, R., Highlander, S. K., Jones, M. B., Sirlin, C. B., Schnabl, B., Brinkac, L., Schork, N., Chen, C. H., Brenner, D. A., Biggs, W., Yooseph, S., … Nelson, K. E. (2017). Gut microbiome–based metagenomic signature for non-invasive detection of advanced fibrosis in human nonalcoholic fatty liver disease. Cell Metabolism, 25(5), 1054–1062.

13. Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J. A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., … Segal, E. (2015). Personalized nutrition by prediction of glycemic responses. Cell, 163(5), 1079–1094.

14. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453.

15. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: addressing ethical challenges. PLoS Medicine, 15(11), e1002689.

16. McMahan, B., Moore, E., Ramage, D., Hampson, S., & Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 1273–1282.

17. Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L. J., Fanzo, J.,

Hawkes, C., Zurayk, R., Rivera, J. A., De Vries, W., Sibanda, L. M., … Murray, C. J. L. (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. The Lancet, 393(10170), 447–492.

18. Zhao K, et al. Phyllostachys nigra (Lodd ex Lindl) derived polysaccharide with enhanced glycolipid metabolism regulation and mice gut microbiome. Int J Biol Macromol. 2024;257:128588.

19. Cani, P. D., & Delzenne, N. M. (2011). The gut microbiome as therapeutic target. Pharmacology & Therapeutics, 130(2), 202–212.

20. Ginsburg, G. S., & Phillips, K. A. (2018). Precision medicine: from science to value. Health Affairs, 37(5), 694–701.