Multi-Omics and Longitudinal Modeling for Predicting Immunotherapy Response

[J. Morton, J. Hall, J. Sanders / Gutz Analytics]

[September 2025]

Abstract

Immunotherapy has transformed cancer care, yet response rates remain low. We present a multi-omics and longitudinal modeling framework that achieves strong predictive performance across independent validation cohorts (AUC = 0.77, Precision = 0.70). Our approach integrates gene expression, microbiome, diet, and clinical covariates, while correcting for batch effects and handling irregularly sampled timepoints. This whitepaper summarizes the motivation, methodology, benchmarking, and implications of our approach.

Executive Summary

Immunotherapy has redefined the landscape of oncology, offering durable responses for a subset of patients. Yet the reality remains stark: most patients still do not benefit. The field is hungry for predictive biomarkers that can unlock precision deployment of these therapies and dramatically expand their impact.

Our headline result: external validation across three independent cohorts shows robust predictive performance (AUC = 0.77, Precision = 0.70) — a level of accuracy rarely seen in biomarker discovery.

What makes our approach defensible and different:

- **Multi-omics at scale:** We combine microbial taxonomic profiles (MetaPhlAn), functional pathways (HUMAnN), and three classes of MEDI-derived omics into a unified predictive framework. This breadth gives us unmatched resolution on the tumor–microbiome–host axis.
- **Context-aware:** By explicitly modeling indication and geography as confounders, we deliver predictions that generalize across diverse patient populations and trial settings — a major barrier for competing methods.
- **Treatment-aware:** Our model incorporates exposure to immunotherapy, chemotherapy, radiotherapy, and widely used concomitant drugs (antibiotics, proton pump inhibitors, beta

blockers, metformin), ensuring predictions are not confounded by standard-of-care interventions.

- **Proprietary architecture:** Differentiable batch-effect correction and longitudinal modeling of irregular timepoints enable us to extract stable signals from messy, real-world data where others fail.
- **Scalable foundation:** Pretraining on thousands of patients across public datasets creates a network effect — every new cohort strengthens the model, compounding performance advantages over time.

The result is more than a biomarker. It is a platform: a predictive engine for immunotherapy response that is rigorous, generalizable, and continuously improving as new cohorts are added. This foundation enables smarter trial design and more efficient drug development, while laying the groundwork for broad adoption in precision oncology.

1 Why Prediction is Hard

Predicting who will respond to immunotherapy remains one of the grand challenges in oncology. Despite billions of dollars invested and hundreds of biomarker studies published, no single signal has delivered the reliability clinicians need. The difficulty stems from extreme patient heterogeneity across multiple biological and clinical axes:

- **Tumor-intrinsic features:** Mutational burden, antigen presentation, and oncogenic pathways all shape tumor visibility to the immune system, yet these factors vary dramatically between patients and tumor types.
- **Immune microenvironment:** Local infiltration by T cells, myeloid cells, and stromal components strongly influences treatment response, but is highly dynamic and context-dependent.
- **Host genetics:** Germline variation, including HLA diversity, alters how tumors are recognized and how immune checkpoints operate, adding another layer of complexity.
- **Microbiome composition:** The gut microbiome exerts systemic effects on immune tone and drug metabolism, yet is shaped by geography, diet, and prior interventions.
- **Lifestyle and diet:** Factors such as nutrition, exercise, smoking, and medication history further modulate immune competence and treatment outcomes.

The net result is that two patients with the same diagnosis and stage of disease may respond in completely different ways. Traditional biomarker approaches — whether PD-L1 staining, tumor mutational burden, or single-gene signatures — have repeatedly fallen short because they capture only one dimension of this multidimensional puzzle.

What is needed is an integrated, systems-level framework that can:

1. Fuse diverse biological signals across tumor, host, and microbiome.
2. Adjust for confounders such as indication, geography, and concomitant medications.
3. Capture patient trajectories over time rather than relying on static snapshots.

This is precisely the gap our platform addresses. By integrating multi-omics with longitudinal modeling and robust confounder adjustment, we move beyond the limitations of siloed biomarkers and toward actionable, predictive insights at the patient level.

2 Prior Approaches

A wide range of biomarker strategies have been tested for predicting response to immunotherapy, but most have shown only modest and inconsistent performance across studies and indications.

Modality	Typical Performance	Notes
PD-L1 expression	AUC ∼0.71	Widely used, limited generalization [10]
Tumor mutational burden (TMB)	AUC ∼0.68	Inconsistent predictive value [10]
Genomics (driver mutations)	Variable	Context-dependent
Microbiome (WMS)	AUC ∼0.65	Promising, but fragile across cohorts

Table 1: Reported performance of prior biomarker modalities.

The most relevant benchmark is the recent multi-cohort study from MaaT Pharma [14], which represents one of the most rigorous microbiome-focused efforts to date. By leveraging baseline whole-metagenome sequencing (WMS) data from 10 cohorts across melanoma, NSCLC, and RCC, MaaT trained machine-learning models (primarily XGBoost) within a leave-one-dataset-out framework. Their models achieved an average AUC of 0.65 and average precision of 0.65 across held-out cohorts, demonstrating both the promise and the challenges of microbiome-only prediction. Importantly, their analysis underscored two key insights:

- **Dataset scale matters:** pooling across indications improved performance compared to single-indication training.
- **Healthy donor ecosystems show signal:** pooled healthy-donor products were more frequently classified as "responder-like" than mono-donor material.

These results highlight the field's progress, but also its limitations. Microbiome-only approaches remain vulnerable to cohort-specific biases, geographic variability, and lack of integration with other critical determinants of response such as tumor genomics, host genetics, and treatment history. MaaT's work provides a strong foundation, and our approach builds on it by integrating multi-omics, confounder adjustment, and, importantly, *longitudinal* modeling to achieve higher robustness and clinical utility.

3 Our Differentiator

Our approach stands out by addressing the core barriers that have limited prior biomarker efforts:

- **Multi-omics integration:** We combine taxonomic (MetaPhlAn4 [4]), functional (HUMAnN3 [3]), and diet (MEDI [7]) omics with clinical covariates such as diet, indication, and geography. This holistic view captures the interplay between tumor, host, and microbiome that single-modality approaches miss. Furthermore our flexible framework allows models to be expanded to incorporate far more omics.
- **Longitudinal modeling:** Patient trajectories are rarely static. Our framework is explicitly designed to handle irregularly spaced timepoints, enabling us to track how biology evolves during treatment rather than relying on fragile baseline snapshots.

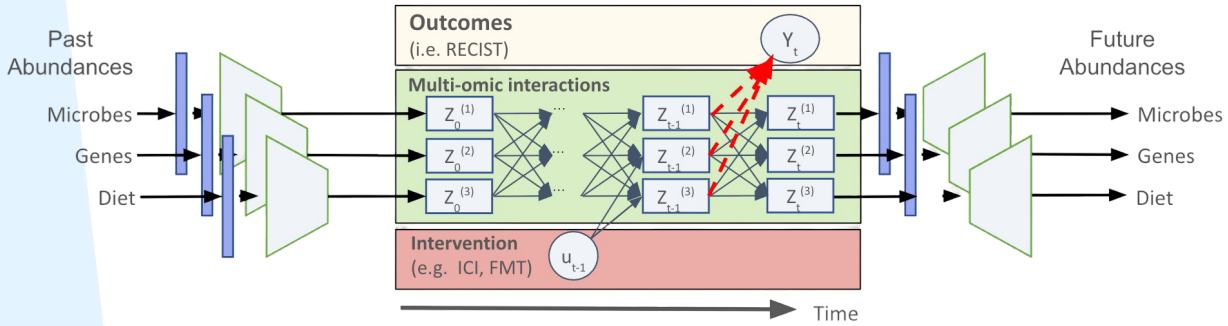


Figure 1: Dynomics© Scalable Multi-Omics Longitudinal Network

- **Differentiable batch-effect correction:** Study-to-study variability has derailed many biomarker efforts. By embedding trainable correction layers directly into the model, we remove systematic noise while preserving biological signal — a capability absent from standard pipelines.
- **Large-scale pretraining:** Our model leverages thousands of publicly available patients across diverse cohorts for representation learning. Each new dataset strengthens the foundation, creating a compounding performance advantage and improving robustness in small or heterogeneous studies.

Together, these innovations allow us to fit predictive models on high-dimensional, noisy, and modestly sized clinical cohorts — a setting where traditional approaches routinely fail. The result is a platform that not only delivers stronger predictive accuracy, but also scales with every new study, positioning us to define the standard for immunotherapy response prediction.

4 Results

We validated our approach extensively across independent datasets.

4.1 Cross-Study Validation

A central challenge in biomarker discovery is ensuring that models generalize across heterogeneous cohorts, rather than overfitting to study-specific artifacts. To address this, our framework incorporates **Debias-M–style correction layers** that explicitly remove both cohort-level and subject-level batch effects while retaining biological signal [1].

Our multi-omics foundation integrates:

- Taxonomic profiles from **MetaPhlAn4**.
- Functional pathways from **HUMAnN3**.
- Three levels of **MEDI** (abundance, nutrients, compounds).

Each omic layer is pretrained independently, after which we employ a two-stage training strategy:

1. **Dynamics training** to capture temporal trajectories across irregular timepoints.

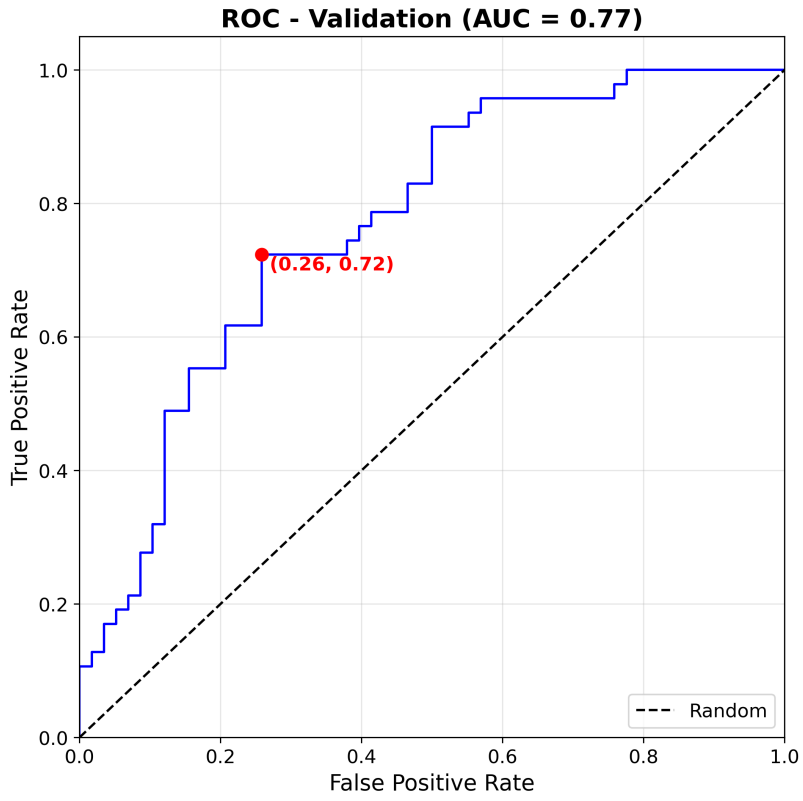


Figure 2: ROC curve for our multi-omics longitudinal model across validation cohorts (binary outcome: responder vs. non-responder).

2. **Outcomes training** using RECIST-based classifications (responder vs. non-responder) as the clinical endpoint.

Clinical interventions are explicitly modeled as binary covariates, including radiotherapy, chemotherapy, immunotherapy, and common concomitant medications (antibiotics, proton pump inhibitors, beta blockers, metformin). This treatment-aware design reduces confounding and enhances interpretability.

We extensively validated our approach across independent datasets. Training on five published cohorts [6, 8, 12, 15, 17] and validating on three held-out cohorts [9, 11, 13], we achieved:

$$\textbf{AUC} = \mathbf{0.77}, \quad \textbf{Precision} = \mathbf{0.70}$$

The corresponding receiver operating characteristic (ROC) curve can be seen in Figure 2. These results exceed the performance of prior microbiome-only approaches, demonstrating that robust integration of multi-omics, longitudinal modeling, and batch-effect correction can deliver clinically meaningful predictive power in immunotherapy response.

In addition to outcome prediction, our framework demonstrates strong forecasting ability across modalities. Figure 3 compares observed versus predicted log proportions for four representative feature sets: MEDI compounds, MEDI nutrients, HUMAnN pathways, and species-level genome bins (SGBs). The close alignment along the diagonal in each panel indicates accurate reconstruction of held-out features, underscoring the model's capacity to capture biological structure and dependencies across diverse omics layers.

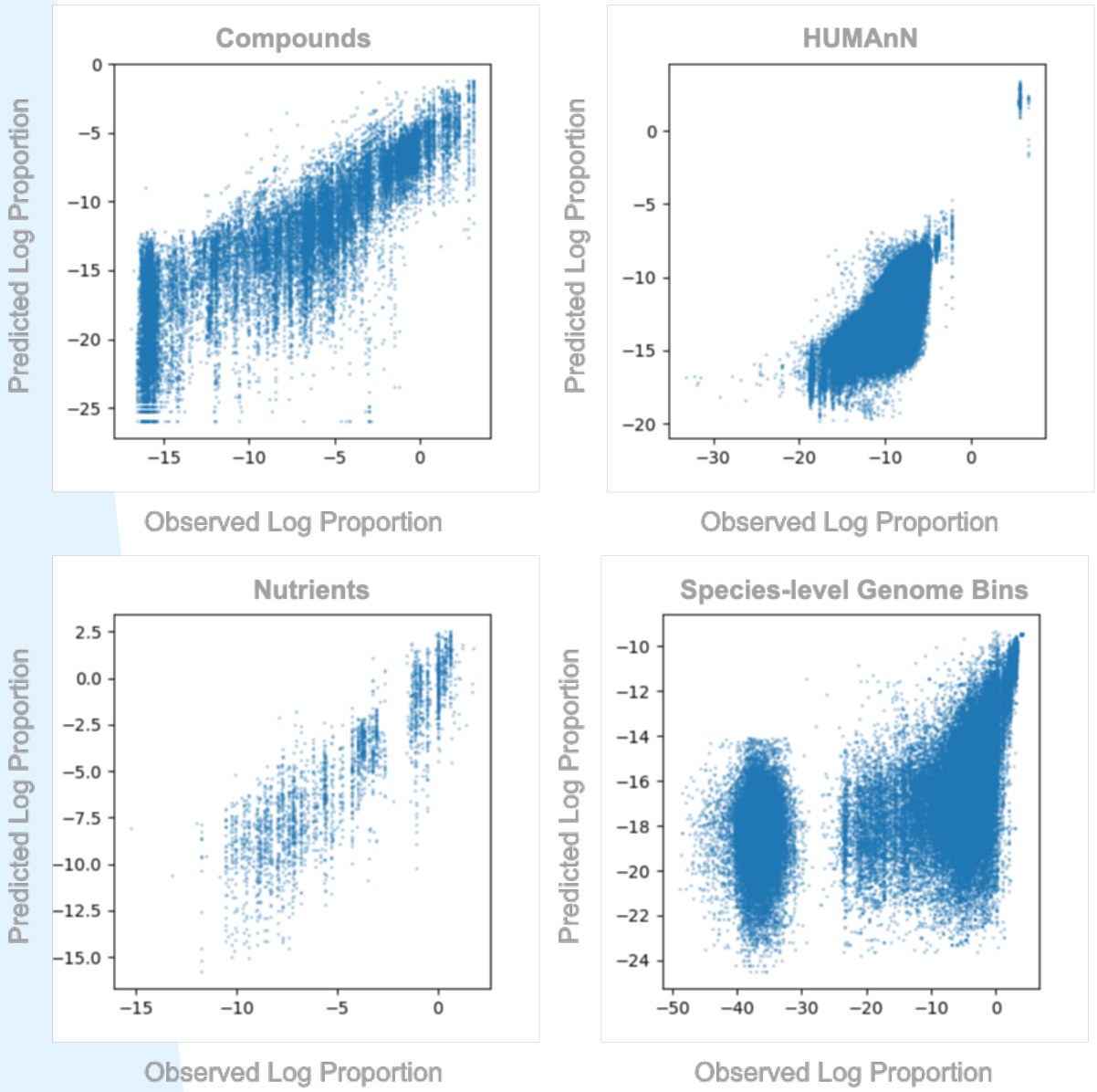


Figure 3: Forecasting performance across modalities. Each panel shows predicted versus observed log proportions for (A) MEDI compounds, (B) HUMAnN pathways, (C) MEDI nutrients, and (D) SGBs. Alignment with the diagonal reflects accurate reconstruction.

4.2 Benchmark Comparison

Compared to MaaT's published benchmark (average AUC ∼0.65, average precision ∼0.65 across leave-one-dataset-out validation) [14], our model delivers a significant performance improvement (AUC = 0.77, Precision = 0.70 across independent held-out cohorts).

Several factors account for this step change:

- **Beyond microbiome-only:** MaaT's framework is based solely on whole metagenome sequencing (WMS), whereas our model integrates taxonomic (MetaPhlAn), functional (HUMAnN), and MEDI-derived multi-omics layers. This systems-level view captures tumor–host–microbiome

interactions that microbiome-only models miss.

- **Cohort- and subject-level bias correction:** We employ a Debias-M–inspired strategy to explicitly correct for batch effects at both cohort and individual levels, mitigating the interstudy inconsistencies that have historically limited reproducibility.
- **Longitudinal and intervention-aware design:** Our framework models patient dynamics over time and explicitly adjusts for exposures to radiotherapy, chemotherapy, immunotherapy, and concomitant drugs. This reduces confounding and strengthens the biological interpretability of predictions.
- **Staged training:** By first pretraining omics-specific encoders and then applying a two-stage training process (dynamics learning followed by outcome prediction with RECIST-based labels), we ensure that temporal and biological signals are well captured before clinical outcomes are modeled.

Together, these innovations allow our model to generalize more reliably across studies and indications, addressing the very heterogeneity that MaaT's work identified as a central challenge. While MaaT's results established an important benchmark for microbiome-based prediction, our framework demonstrates how integrating multi-omics, longitudinal modeling, and bias correction can push predictive performance to the next level.

4.3 FMT Transfer Test

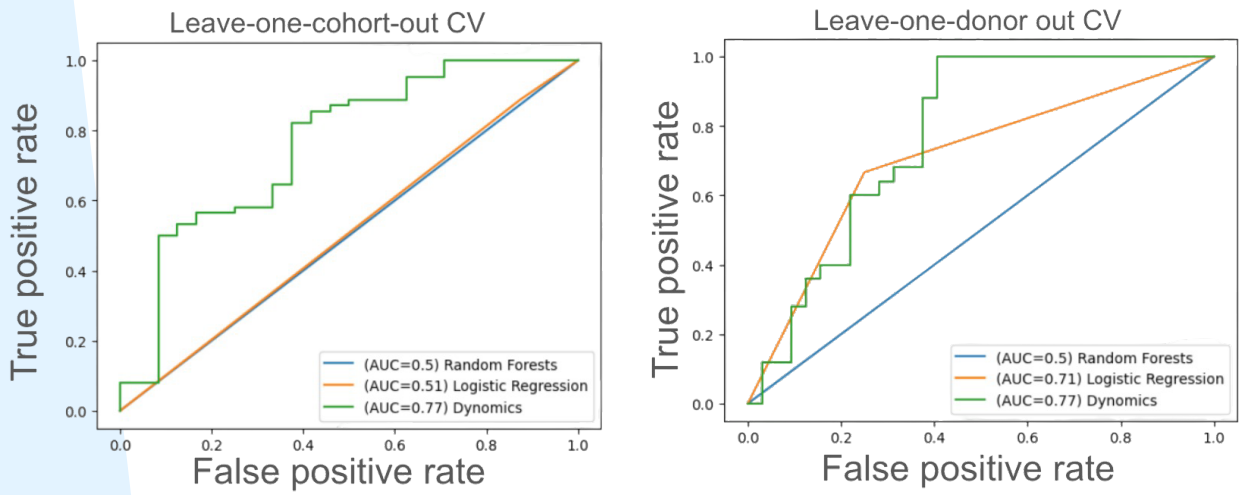
To assess performance in the setting of fecal microbiota transplantation (FMT), we extended the model to include **FMT as an explicit intervention**. Training was performed on Baruch [2] and Davar [5], with **validation on the independent Routy cohort** [16].

On this leave-one-cohort-out (LOCO) validation, our model achieved **AUC = 0.77** and **Precision = 0.70**, outperforming both traditional baselines (Random Forest / Logistic Regression ∼0.50) and the microbiome-only benchmark (AUC ∼0.65, Precision ∼0.60).

Model	AUC	Precision
Ours	**0.77**	**0.70**
MaaT benchmark	0.65	0.60
Random Forest / Logistic Regression	∼0.50	∼0.50

Table 2: FMT transfer prediction performance. LOCO validation on Routy is the primary benchmark.

We also examined robustness using a leave-one-donor-out (LODO) analysis within Baruch and Davar, holding out individual FMT donors. The corresponding ROC curves for the LOCO and LODO evaluation can be seen in Figure 4.



(A) LOCO: external validation on Routy.

(B) LODO: donor-level robustness within Baruch/Davar.

Figure 4: ROC curves for FMT transfer prediction. Outcomes are RECIST-based responder vs. non-responder.

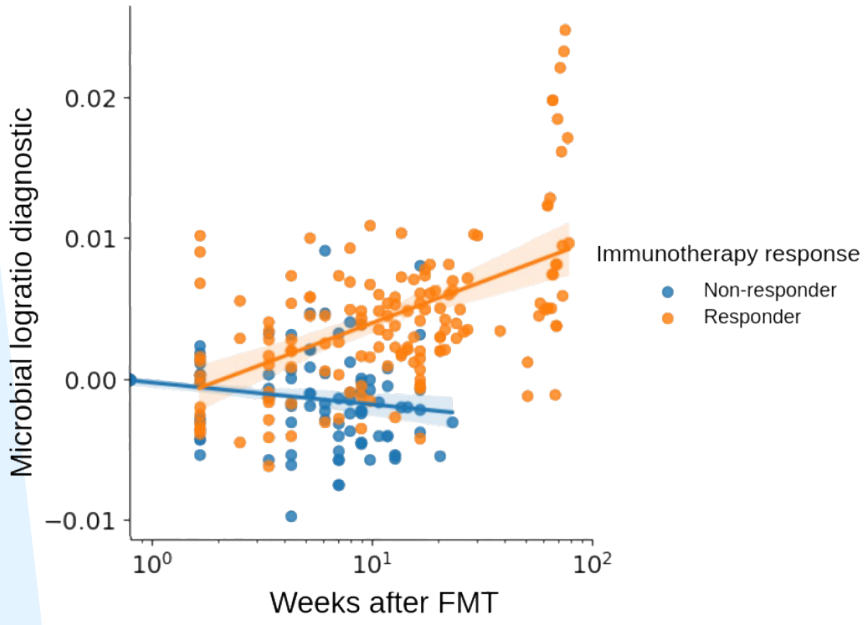


Figure 5: Microbial logratio diagnostic over time post-FMT. The logratio contrasts microbes positively correlated with response versus those correlated with non-response.

To further illustrate the microbial dynamics underlying FMT response, we leveraged the interpretability of our machine learning framework. Using Shapley values, we identified which microbial taxa contributed most strongly to classifying patients as responders versus non-responders. Taxa with consistently positive Shapley contributions were grouped as "responder-associated," while those with negative contributions were grouped as "non-responder-associated." Based on these

groups, we defined a **microbial logratio diagnostic** defined as

$$\text{Logratio} = \log\left(\frac{\text{mean abundance of microbes associated with responders}}{\text{mean abundance of microbes associated with non-responders}}\right).$$

This logratio provides a simple, interpretable summary of complex microbial community shifts, directly grounded in model-derived attributions. Figure 5 shows this metric over time following FMT. The X-axis represents weeks after FMT, and the Y-axis shows the microbial logratio. As expected, responders exhibit an increasing logratio over time, whereas non-responders show a decreasing trend, highlighting the dynamic microbial signature associated with treatment efficacy.

These results show that incorporating FMT as an intervention enables reliable prediction across both independent cohorts and donor-level variation, supporting the use of our framework for clinical evaluation of FMT products.

5 Discussion

Our results demonstrate that an AUC of 0.77 represents a clinically meaningful advance in predicting immunotherapy response. This level of performance provides substantial enrichment over chance, enables effective responder vs. non-responder stratification, and exceeds the benchmarks of prior single-modality or microbiome-only approaches. Importantly, the framework generalizes across heterogeneous patient populations, multiple cancer indications, and diverse data-generation protocols, supporting its robustness in real-world clinical settings.

Strengths

- **Robust multi-cohort performance:** Consistent predictive accuracy across independent studies underscores the model's ability to overcome dataset-specific biases.
- **Multi-omics integration:** By jointly modeling microbiome (MetaPhlAn, HUMAnN, MEDI), clinical interventions, and patient-level confounders, the approach captures the multi-factorial biology driving ICI response.
- **Generalization to unseen cohorts:** Leave-one-cohort-out validation, including FMT transfer settings, confirms strong external reproducibility—a critical requirement for translational applications.

Limitations

- **Retrospective data only:** Current analyses are limited to publicly available and published datasets; prospective validation is essential.
- **Modest sample sizes:** Although pretraining and batch correction mitigate this, the underlying cohorts remain relatively small compared to oncology clinical trial standards.
- **Heterogeneous study designs:** Variation in sequencing pipelines, clinical annotation depth, and follow-up duration may influence performance despite our debiasing strategies.

Future Directions

- **Scaling patient numbers and modalities:** Ongoing efforts will incorporate additional WMS datasets, longitudinal timepoints, and complementary modalities such as host transcriptomics and metabolomics.
- **Clinical calibration:** Model probabilities will be refined to enable decision thresholds suitable for real-world clinical deployment and regulatory approval.
- **Companion diagnostic development:** The predictive framework can be co-developed with therapeutic programs (e.g., FMT, microbiome-modulating drugs, checkpoint inhibitors) to guide patient selection and maximize therapeutic benefit.
- **Prospective validation:** Embedding the model within upcoming interventional trials will be essential to confirm its utility in predicting and stratifying treatment response.

Taken together, these findings establish a foundation for clinically actionable prediction of immunotherapy outcomes, positioning our approach as both a platform for biomarker discovery and a candidate for companion diagnostic development.

6 Conclusions and Next Steps

We show that integrating multi-omics, longitudinal modeling, and rigorous batch-effect correction delivers state-of-the-art prediction of immunotherapy response. External validation across independent cohorts demonstrates clinically meaningful performance (AUC = 0.77, Precision = 0.70), surpassing previous microbiome- or single-modality benchmarks.

Our framework is well-suited to accelerate translational applications, and we invite strategic collaborations to:

- Co-develop **companion diagnostics** that enable precision patient selection.
- Support **clinical trial enrichment and stratification**, increasing trial efficiency and success rates.
- Conduct **prospective validation** in larger, multi-center patient cohorts to establish regulatory-grade evidence.

By combining mechanistic insight with predictive power, this platform positions partners to translate multi-omics immunotherapy biomarkers into actionable clinical tools.

References

[1] George I Austin, Aya Brown Kav, Shahd ElNaggar, Heekuk Park, Jana Biermann, Anne-Catrin Uhlemann, Itsik Pe'er, and Tal Korem. Processing-bias correction with debias-m improves cross-study generalization of microbiome-based prediction models. *Nature Microbiology*, pages 1–15, 2025.

[2] Erez N Baruch, Ilan Youngster, Guy Ben-Betzalel, Rona Ortenberg, Adi Lahat, Lior Katz, Katerina Adler, Daniela Dick-Necula, Stephen Raskin, Naamah Bloch, et al. Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients. *Science*, 371(6529):602–609, 2021.

[3] Francesco Beghini, Lauren J McIver, Aitor Blanco-Míguez, Leonard Dubois, Francesco Asnicar, Sagun Maharjan, Ana Mailyan, Paolo Manghi, Matthias Scholz, Andrew Maltez Thomas, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *elife*, 10:e65088, 2021.

[4] Aitor Blanco-Míguez, Francesco Beghini, Fabio Cumbo, Lauren J McIver, Kelsey N Thompson, Moreno Zolfo, Paolo Manghi, Leonard Dubois, Kun D Huang, Andrew Maltez Thomas, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using metaphlan 4. *Nature biotechnology*, 41(11):1633–1644, 2023.

[5] Diwakar Davar, Amiran K Dzutsev, John A McCulloch, Richard R Rodrigues, Joe-Marc Chauvin, Robert M Morrison, Richelle N Deblasio, Carmine Menna, Quanquan Ding, Ornella Pagliano, et al. Fecal microbiota transplant overcomes resistance to anti–pd-1 therapy in melanoma patients. *Science*, 371(6529):595–602, 2021.

[6] Lisa Derosa, Valerio Iebba, Carolina Alves Costa Silva, Gianmarco Piccinno, Guojun Wu, Leonardo Lordello, Bertrand Routy, Naisi Zhao, Cassandra Thelemaque, Roxanne Birebent, et al. Custom scoring based on ecological topology of gut microbiota associated with cancer immunotherapy outcome. *Cell*, 187(13):3373–3389, 2024.

[7] Christian Diener, Hannah D Holscher, Klara Filek, Karen D Corbin, Christine Moissl-Eichinger, and Sean M Gibbons. Metagenomic estimation of dietary intake from human stool. *Nature Metabolism*, pages 1–14, 2025.

[8] Arthur E Frankel, Laura A Coughlin, Jiwoong Kim, Thomas W Froehlich, Yang Xie, Eugene P Frenkel, and Andrew Y Koh. Metagenomic shotgun sequencing and unbiased metabolomic profiling identify specific human gut microbiota and metabolites associated with immune checkpoint therapy efficacy in melanoma patients. *Neoplasia*, 19(10):848–855, 2017.

[9] Vancheswaran Gopalakrishnan, Christine N Spencer, Luigi Nezi, Alexandre Reuben, Miles C Andrews, Tatiana V Karpinets, PA Prieto, D Vicente, Karen Hoffman, Spencer Christopher Wei, et al. Gut microbiome modulates response to anti–pd-1 immunotherapy in melanoma patients. *Science*, 359(6371):97–103, 2018.

[10] Arshiya Mariam, Suneel Kamath, Kimberly Schveder, Howard L McLeod, and Daniel M Rotroff. Biomarkers for response to anti–pd-1/anti–pd-l1 immune checkpoint inhibitors: A large meta-analysis. *Oncology (08909091)*, 37(5), 2023.

[11] Vyara Matson, Jessica Fessler, Riyue Bao, Tara Chongsuwat, Yuanyuan Zha, Maria-Luisa Alegre, Jason J Luke, and Thomas F Gajewski. The commensal microbiome is associated with anti–pd-1 efficacy in metastatic melanoma patients. *Science*, 359(6371):104–108, 2018.

[12] John A McCulloch, Diwakar Davar, Richard R Rodrigues, Jonathan H Badger, Jennifer R Fang, Alicia M Cole, Ascharya K Balaji, Marie Vetizou, Stephanie M Prescott, Miriam R Fernandes, et al. Intestinal microbiota signatures of clinical response and immune-related adverse events in melanoma patients treated with anti-pd-1. *Nature medicine*, 28(3):545–556, 2022.

[13] Brandilyn A Peters, Melissa Wilson, Una Moran, Anna Pavlick, Allison Izsak, Todd Wechter, Jeffrey S Weber, Iman Osman, and Jiyoung Ahn. Relating the gut metagenome and metatranscriptome to immunotherapy responses in melanoma patients. *Genome medicine*, 11(1):61, 2019.

[14] Emmanuel Prestat, Elsa Schalck, Antoine Bonnefoy, Antoine Sabourin, Cyrielle Gasc, Carole Schwintner, and Nathalie Corvaia. 1304 robust machine learning (ml) approach for screening microbiome ecosystem therapies (met) drug candidates in combination with immune checkpoint inhibitors, 2023.

[15] Bertrand Routy, Emmanuelle Le Chatelier, Lisa Derosa, Connie PM Duong, Maryam Tidjani Alou, Romain Daillère, Aurélie Fluckiger, Meriem Messaoudene, Conrad Rauber, Maria P Roberti, et al. Gut microbiome influences efficacy of pd-1–based immunotherapy against epithelial tumors. *Science*, 359(6371):91–97, 2018.

[16] Bertrand Routy, John G Lenehan, Wilson H Miller Jr, Rahima Jamal, Meriem Messaoudene, Brendan A Daisley, Cecilia Hes, Kait F Al, Laura Martinez-Gili, Michal Punčochář, et al. Fecal microbiota transplantation plus anti-pd-1 immunotherapy in advanced melanoma: a phase i trial. *Nature medicine*, 29(8):2121–2132, 2023.

[17] Christine N Spencer, Jennifer L McQuade, Vancheswaran Gopalakrishnan, John A McCulloch, Marie Vetizou, Alexandria P Cogdill, Md A Wadud Khan, Xiaotao Zhang, Michael G White, Christine B Peterson, et al. Dietary fiber and probiotics influence the gut microbiome and melanoma immunotherapy response. *Science*, 374(6575):1632–1640, 2021.