AI classifier integrating taxonomic and functional microbial features in melanoma and NSCLC improves prediction of response to ICI

ESMO Congress 2026 | AI for clinical research and drug development

FPN : 37P

Joël Doré[1], Liyang Diao[2], Elsa Schalck[3], Germain Pinganaud[4], Antoine Sabourin[4], Grégoire Noiret[4], Cyrielle Gasc[4], Lisa Derosa[5,6], Laurence Zitvogel[5], Anne-Claire Toffart[7], Dalil Hannani[8], Antoine Bonnefoy[3], Sheri Simmons[4], Emmanuel Prestat[4]

[1] INRAE, French National Research Institute for Agriculture, Food and Environment, France · [2] Sprout Informatics · [3] Euranova, Marseille, France · [4] MaaT Pharma · [5] Gustave Roussy (GRCC), ClinicObiome, Inserm UMR1367, Microbiota and Mucosal Immunity for Cancer Immunotherapy, Université Paris-Saclay, Villejuif, France · [6] Department of Medical Oncology, Gustave Roussy, Villejuif, France · [7] Thoracic Oncology Unit, Pulmonology, Grenoble University Hospital, Grenoble, France · [8] Univ. Grenoble Alpes, CNRS, UMR 5525, VetAgro Sup, Grenoble INP, TIMC, Grenoble, France

TAKEAWAYS

1. MaaT Pharma's gutPrint® platform developed **an AI framework to classify microbiome samples as "responder-like"** based on baseline stool metagenomic profiles from ICI-treated patients, thereby supporting the evaluation of microbiome-based therapeutic approaches in melanoma, NSCLC and RCC.
2. **Performance improved** when taxonomic and functional features were integrated and **multiple datasets and indications** were combined for model training.
3. The most promising published AUCs appear to reflect overfitting: **no model has reproduced these performances across independent cohorts or indications**.

   We encourage the use of more **robust validation strategies**, such as Leave-One-Dataset-Out cross-validation, even when they result in less spectacular but more realistic AUCs.

BACKGROUND

- MaaT Pharma, a clinical-stage biotech pioneer in the development of Microbiome Ecosystem Therapies in oncology, is defining its strategy to customize products in immuno-oncology related indications.
- Immune checkpoint inhibitors (ICIs) have improved outcomes in melanoma and non-small cell lung cancer (NSCLC), with multiple studies demonstrating a clear association between gut microbiome composition and ICI efficacy (Figure 1). In addition, FMT proofs of concept were performed [1, 2, 3, 4, 5, 6].
- Baseline gut microbiome features may predict ICI response. However, inter-cohort heterogeneity limits robust discovery and prevents the validation of robust and clinically performant predictive models.
- Baseline stool shotgun metagenomic datasets from ICI treated cancer patients were gathered along with their response as defined by Response Evaluation Criteria in Solid Tumors (RECIST) and provided by each source study.
- Interstudy inconsistencies (Figures 1 and 2) were observed in microbiome signature findings [7].
- Resolving this heterogeneity is essential to isolating reproducible response signals and developing a reliable screening algorithm for microbiota-based drug candidates.

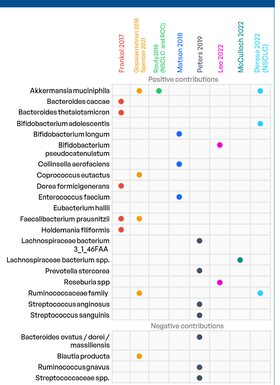


Figure 1: Comparison of species identified in the literature. Only species mentioned in the abstracts of the related papers were gathered.

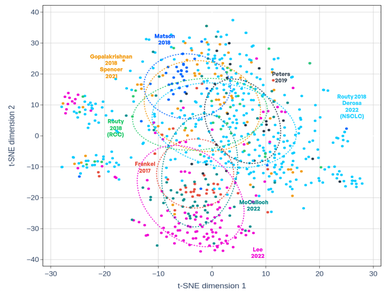


Figure 2: t-SNE of baseline samples from 10 public cohorts based on genus, KEGG Pathways and alpha diversity. Preprocessing steps: relative abundance normalization and standard scaling.

METHODS

As part of gutPrint®, MaaT Pharma's bioinformatics and AI platform :

- Baseline whole-metagenome sequencing (WMS) profiles were gathered across 16 published cohorts encompassing melanoma (N = 536, 11 cohorts), NSCLC (N = 425, 4 cohorts), and RCC (N = 66, 1 cohort) (Figure 3). Raw data were uniformly processed via gutPrint® MgRunner to generate taxonomic and functional profiles. Downstream analyses included only successfully processed samples with annotated clinical outcomes (Table 1).
- For NSCLC and melanoma indications, taxon-response associations were estimated separately across cohorts using models adjusted for antibiotic exposure. Evidence for positive and negative associations was aggregated separately using a modified Fisher combination test. A cross-study directional consistency score, defined as the ratio of the variance of absolute effect estimates to the variance of signed effect estimates, was used to quantify whether cohort-specific effects consistently pointed in the same direction. The directional association score was then calculated as the difference between the aggregated positive and negative scores and compared between melanoma and NSCLC for genera shared across both indications.
- A robust machine learning (ML) approach for screening microbiome samples and predicting potential responder status to ICIs was developed using a combination of taxonomic and functional features. We relied on the Area Under the ROC Curve (AUC) and precision to assess cross-cohort predictions performances computed through a **Leave-One-Dataset-Out** validation strategy.

1027 patients from **16 publicly available cohorts of patients treated with ICI(s)** for which **shotgun sequencing** of stools microbiota were performed.



Melanoma N= 536



Renal Cell Carcinoma N= 66



Non-Small Cell Lung Cancer N= 425

Figure 3: Summary of baseline Whole Metagenome Sequencing (WMS) public datasets.

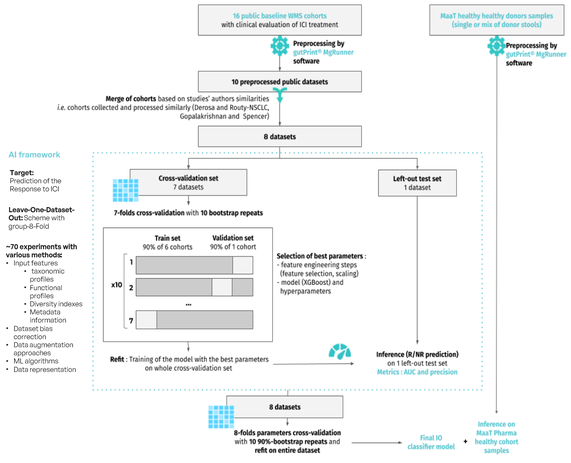
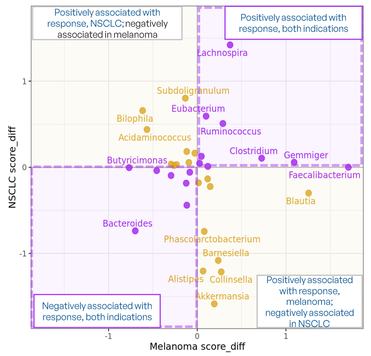


Figure 4: Data processing and AI framework methodologies. All samples were processed by gutPrint® MgRunner to get taxonomic and functional features. The processed datasets were distributed sequentially in the cross-validation set used to select parameters (and then used as the training set), and in the left-out test dataset. This strategy was repeated for each dataset as left-out, and for the 70 designed experiments.

RESULTS

1 Directional response-association signal: melanoma vs. NSCLC (genus level)

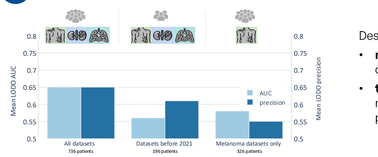
Score_diff = combined_score_pos - combined_score_neg (ABX in model)



- Separate analyses of melanoma and NSCLC cohorts identified a limited number of statistically significant signals, including the genera Lachnospira and Faecalibacillus, which were positively associated with response. Blautia was also positively associated with response in melanoma.
- Given high taxonomic variability, we hypothesized that functional profiles could be more robust.

Figure 5: Genera with consistently positive and negative association in both melanoma and NSCLC
Top positively associated genus in both: Lachnospira
Top negatively associated genus in both: Bacteroides

2 Performance by dataset and indication



Despite the diverse data sources and indications :

- **more datasets** (since 2021) improved the classification performances
- **the multi-indication** approach surpassed the mono-indication (melanoma) training approach for predictions related to melanoma patients.

Figure 6: Best model's results on multiple dataset combinations

	Routy 2018 [8] → Derosa 2022 (NSCLC) [9]	Routy 2018 (RCC) [8]	Frankel 2017 [10]	Gopala-krishnan 2018 [11] → Spencer 2021 [12]	Lee 2022 [13]	McCulloch 2022 [14]	Matson 2018 [15]	Peters 2019 [18]	ALL MELANOMA	ALL DATASETS
Indication										NSCLC RCC Melanoma
ICI	Anti-PD1	Anti-PD1	Anti-PD1 / Anti-CTLA4	Anti-PD1	Anti-PD1 / Anti-CTLA4	Anti-PD1	Anti-PD1	Anti-PD1 / Anti-CTLA4	Anti-PD1 / Anti-CTLA4	Anti-PD1 / Anti-CTLA4
# patients	381	29	32	57	121	58	36	22	326	736
Proportions of responders	51.7%	62.1%	56.3%	52.6%	52.9%	53.5%	36.1%	50%	51.2%	51.9%
AUC	0.57	0.71	0.66	0.65	0.66	0.52	0.66	0.74	Mean: 0.65	Mean: 0.65
Precision	0.56	0.81	0.68	0.63	0.66	0.55	0.57	0.67	Mean: 0.63	Mean: 0.65

Table 1: Best model's performance by dataset and indication

- The best performing experiment provided models based on genera, KEGG Pathways and alpha diversity features as inputs treated with the XGBoost algorithm. AUCs range from 0.52 to 0.74 depending on the left-out cohort (average **AUC** = 0.65) and a precision that ranges between 0.55 and 0.81 (average **precision** = 0.65).
- Those results **exceeded** those reported by a melanoma-centered study with a comparable assessment method [13] for common cohorts (where Matson AUC = 0.61, Gopalakrishnan AUC = 0.56, Frankel AUC = 0.63, and mean Lee datasets AUC = 0.60).

3 Performance by dataset and indication

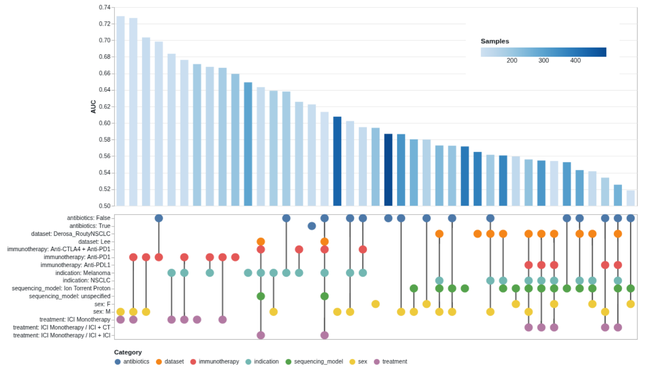


Figure 7: Model performance across subsets
The UpSet plot shows the mean AUC for intersections of clinical characteristics. Bars are ranked by decreasing AUC, and blue intensity indicates subset size. For example, the 4th bar shows an AUC of almost 0.70 for patients who did not receive antibiotics and were treated with anti-PD1.
Only groups with minimum 100 samples are displayed.

- IO Classifier performance varied across clinical subgroups. The highest mean AUCs, exceeding 0.70 in selected subsets, were mainly observed in patients receiving anti-PD1 monotherapy and in subsets including melanoma or multiple indications.
- NSCLC-only subsets generally showed lower performance, while the influence of antibiotic exposure remained inconclusive.

CONCLUSIONS

- This study demonstrates the feasibility of predicting ICI response from baseline gut metagenomic profiles across heterogeneous melanoma, NSCLC and RCC cohorts.
- By combining taxonomic and functional features within a Leave-One-Dataset-Out framework, the approach achieved reproducible cross-cohort discrimination and outperformed previously reported results on several publicly available datasets.

Together with clinical proof-of-concept studies showing that microbiome modulation can improve ICI outcomes, these findings support the rationale for developing microbiome-based therapies with specifically designed taxonomic and functional compositions.

References:

[1] Davar D. et al., Science. 2021;371(6529):595–602.
[2] Baruch E.N. et al., Science. 2021;371(6529):602–609.
[3] Routy B. et al., Nat Med. 2023;29(8):2121–2132.
[4] Porcari S. et al., Nat Med. 2026;32(4):1316–1324.
[5] Fernandes R. et al., Nat Med. 2026;32(4):1325–1336.
[6] Duttagupta S. et al., Nat Med. 2026;32(4):1337–1350.
[7] Wojciechowski S. et al., Int J Med Microbiol. 2022;312(7):151560.
[8] Routy B. et al., Science. 2018;359(6371):91–97.
[9] Derosa L. et al., Nat Med. 2022;28:315–324.
[10] Frankel A.E. et al., Neoplasia. 2017;19:848–855.
[11] Gopalakrishnan V. et al., Science. 2018;359(6371):97–103.
[12] Spencer C.N. et al., Science. 2021;374(6575):1632–1640.
[13] Lee K.A. et al., Nat Med. 2022;28:535–544.
[14] McCulloch J.A. et al., Nat Med. 2022;28:545–556.
[15] Matson V. et al., Science. 2018;359(6371):104–108.
[16] Peters B.A. et al., Genome Med. 2019;11(1):61.

Disclosure: JD is co-founder and scientific advisor of MaaT Pharma, GMT Science, Novobiome. He has provided consultancy to Nestle and Danone.