AI-based supervised treatment response prediction from tumor transcriptomics: A large-scale pan-cancer study

Lipika R. Pal[1,4,*], E. Michael Gertz[1,4], Nishanth Ulhas Nair[1,4], Sumit Mukherjee[1,5], Sumeet Patiyal[1,5], Thomas Cantore[1], Emma M. Campagnolo[1], Tian-Gen Chang[1], Saugato Rahman Dhruba[1], Yewon Kim[1], Eldad D. Shulman[1], Padma Sheila Rajagopal[1,2], Danh-Tai Hoang[1], Alejandro A. Schäffer[1,*], Eytan Ruppin[1,3*]

[1] Cancer Data Science Laboratory (CDSL), Center for Cancer Research (CCR), National Cancer Institute (NCI), National Institutes of Health (NIH), Bethesda, MD, USA.

[2] Women's Malignancies Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA.

[3] Jim and Eleanor Randall Department of Surgery, Cedars-Sinai Medical Center, Los Angeles, CA, USA

[4] These authors contributed equally to this work as co-first authors.

[5] These authors contributed equally to this work.

* Corresponding authors:

Lipika R. Pal: lipika.ray@nih.gov

Alejandro A. Schäffer: alejandro.schaffer@nih.gov

Eytan Ruppin: Eytan.Ruppin@cshs.org

ABSTRACT

Precision oncology aims to guide treatment decisions using biomarkers. While DNA-based panels are increasingly applied, RNA transcriptomics remain underused due to limited datasets and the absence of robust models. We assembled the largest transcriptomic resource for drug response prediction to date, spanning 69 cohorts, 3,729 patients, nine cancer types, and six frontline therapies: anti-PD-1/PD-L1 immune-checkpoint inhibitors, trastuzumab, bevacizumab, BRAF inhibitors, paclitaxel, and FAC/FEC (Fluorouracil-Adriamycin-Cyclophosphamide/Fluorouracil-Epirubicin-Cyclophosphamide) chemotherapy. We developed EXPRESSO (EXpression-Profile-RESponSe-Optimizer), a supervised AI-based framework that predicts treatment response from pre-treatment transcriptomes by integrating drug targets and context-specific biomarkers. EXPRESSO achieves ROC-AUCs of 0.64–0.73 and odds ratios of 2.4–4.6 across therapies, outperforming 20 published transcriptomic signatures. Robustness analysis reveals that predictive performance plateaued for some therapies with increasing training cohorts but continued to improve for others. These findings suggest inherent limits of supervised brute-force learning for certain treatments, but additional data and deeper mechanistic modeling may further enhance transcriptomics-based predictors.

INTRODUCTION

Despite major advances in cancer therapeutics, not all patients benefit equally. Response rates vary widely, even within the same cancer type, due to underlying molecular heterogeneity. This challenge has fueled precision oncology, which aims to tailor therapies to individual patients using molecular biomarkers[1]. To date, clinical implementation has largely focused on genomic profiling—particularly targeted DNA sequencing panels—to identify actionable mutations or fusions[2,3]. However, since RNA alterations can reveal information beyond DNA[4], we set out to explore how genome-wide RNA expression can be leveraged to predict patient response to widely used cancer treatments on a large-scale.

Modern oncology now encompasses a broad array of therapeutic modalities, each targeting distinct aspects of tumor biology. These include immune checkpoint blockade, monoclonal antibodies (mAbs), targeted small molecule inhibitors, and traditional cytotoxic chemotherapies[5–8]. Immune checkpoint inhibitors reinvigorate anti-tumor T cell responses by blocking inhibitory pathways[9,10]. Monoclonal antibodies offer high specificity by binding to target antigens on cancer cells or the tumor microenvironment, often engaging immune-mediated mechanisms[11–13]. Targeted small molecules inhibit key signaling proteins—typically kinases—associated with tumor-specific genetic alterations[14,15]. Cytotoxic chemotherapies remain a mainstay for many cancers, exerting their effects through DNA damage or disruption of mitosis[16–18].

As new drugs continue to expand the therapeutic arsenal, the problem of which treatment regimen a cancer patient should receive has grown increasingly complex. Targeted tumor DNA sequencing has been transformative for subsets of patients (e.g., *EGFR* mutations in lung cancer, *BRAF V600E* in melanoma). Yet, many tumors lack actionable alterations, and identified mutations often fail to predict benefit across diverse contexts[19–24]. Moreover, DNA-based features alone do not fully capture dynamic processes such as transcriptional adaptation, tumor–microenvironment interactions, or immune activity[25]. These limitations have driven growing interest in RNA transcriptomics, which provides a functional, real-time readout of tumor biology through global gene expression patterns—capturing coordinated pathway shifts, cell states, and immune responses that may underlie treatment sensitivity[26,27].

Currently, most FDA approved cancer biomarkers for such therapies are genomic (e.g., EGFR, BRAF, BRCA1/2) or proteomic (e.g., HER2 amplification, PD-L1 protein expression)[28–30]. For the 116 drugs with biomarkers listed on the OncoKB site (https://www.oncokb.org/oncology-therapies)[31,32], all biomarkers can be assessed by either DNA sequencing or protein assays[33,34]. Although in some instances, RNA assays substitute for protein measurements, there are currently no AI-based mRNA gene signature biomarkers that integrate the expression levels of multiple genes simultaneously. In immunotherapy, high PD-L1 protein expression and genomic markers such as high tumor mutational burden (TMB) and microsatellite instability (MSI) are routinely used[35–40]. However, these markers often lack pan-cancer applicability and may overlook broader transcriptomic contexts.

The best known FDA-approved prognostic transcriptomic tool is Oncotype DX, a multigene expression assay used in early-stage breast cancer to assess recurrence risk and to guide chemotherapy decisions[41,42]. More broadly, transcriptomic predictors have demonstrated promising results across various therapeutic classes. T-cell–inflamed gene expression profiles, for example, have been associated with response to immune checkpoint inhibitors[43]. Large-scale efforts, such as The Cancer Genome Atlas (TCGA), have further shown that expression-based models can, in certain contexts, outperform DNA-based features[27].

A substantial body of work has attempted to harness transcriptomic data through machine learning to build predictive models of drug response[44–49]. Given the limited availability of large clinical datasets, many efforts have focused on training models using cancer cell lines and transferring models to make predictions to patient tumors. Pioneering studies predicted responses to agents like docetaxel, bortezomib, and erlotinib using cell line-derived gene expression models[50]. Subsequent work extended this strategy by imputing drug response in patient tumors based on *in vitro* transcriptomic models[45,47]. More recently, the Celligner algorithm mapped tumors to mRNA-expression-similar cell lines, revealing transcriptomic alignment often diverged from mutation similarity or methylation similarity[49].

Advanced deep learning strategies[51], including few-shot learning[52], context-aware models (CODE-AE)[53], interpretable frameworks (DrugCell)[54], integrative multi-omics approaches[55], and foundation model using concept-bottleneck architecture[56] have further expanded the landscape of

transcriptome-informed prediction. Several efforts have also focused on combination therapy, using concepts like independent drug action (IDA) to predict synergistic effects from monotherapy data[57]. Despite these innovations, none of these approaches for using transcriptomics to predict patient response has yet been translated into clinical practice.

Parallel to these computational advances, prospective clinical studies have begun to incorporate transcriptomic data into real-world decision-making. For example, Beaubier et. al,[58] reported that RNA-seq and non-RNA immunotherapy-related biomarkers increased the proportion of patients matched to precision therapies from 29.6% to 43.4%. The WINTHER trial demonstrated the utility of transcriptomic-guided therapy selection across diverse cancers[59–63]. Other studies, such as those by[64,65] and the TuPro initiative[66], have shown that multi-omics profiling, including RNA expression, can meaningfully influence clinical outcomes when integrated into tumor board recommendations.

Nevertheless, two major impediments continue to limit the clinical use of transcriptomic data: small sample sizes and poor harmonization between cross-platform expression data[67–69]. With the emergence of expansive transcriptomic datasets—including expression profiles from responders and non-responders across a range of cancers—it is now feasible to systematically evaluate transcriptome-informed predictors[70–73].

In this study, we present a large-scale evaluation of the potential for pre-treatment transcriptomic data to predict patient response across a range of commonly used cancer therapies. To our knowledge, this represents the largest compilation of clinical trial-based transcriptomic datasets assembled for drug response prediction, enabling a systematic investigation of the predictive value of pre-treatment gene expression profiles. We developed EXPRESSO (EXpression Profile RESponSe Optimizer), a biologically-grounded, AI-based, supervised learning framework tailored to specific drugs and cohort settings. In two usage modes, EXPRESSO incorporates either the leading drug's target gene(s) alone or in combination with selected biomarkers derived from differentially expressed genes between responders and non-responders. This framework serves to study the feasibility of harnessing bulk transcriptomic data to inform therapeutic decision-making.

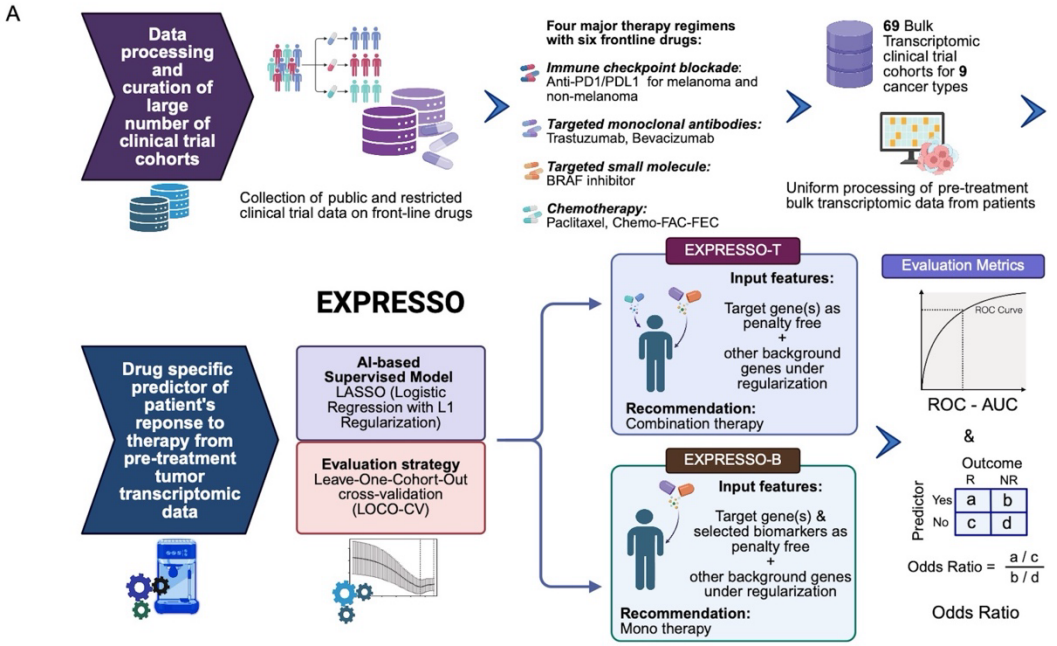
RESULTS

Study overview

To assess the predictive power of bulk transcriptomic data in cancer therapy response, we aggregated publicly available (some requiring dbGaP permission) cancer clinical trial cohorts. In total, we collected 69 pre-treatment bulk transcriptomic cohorts, comprising 3,729 patients across nine cancer types and spanning four major therapeutic modalities: immune-checkpoint-blockade, monoclonal antibody targeted therapy, small molecule targeted therapy, and chemotherapy (**Figure 1A, Supplementary Table S1**). These cohorts included patients treated with drugs of six types: anti-PD-1 checkpoint blockade (targeting the PD-1/PD-L1 axis; for melanoma 9 cohorts, 336 samples and for non-melanoma 14 cohorts, 607 samples), trastuzumab (anti-HER2 monoclonal antibody; 13 cohorts, 494 samples), bevacizumab (8 cohorts, 377 samples), BRAF inhibitors (8 cohorts, 193 samples), paclitaxel (9 cohorts, 933 samples), and chemo-FAC-FEC (Fluorouracil-Adriamycin-Cyclophosphamide / Fluorouracil-Epirubicin-Cyclophosphamide) (9 cohorts, 790 samples) (**Figure 1B**). While most of the patients received these drugs in combination with varying chemotherapies and other drugs (**Figure 1C**, **Supplementary Table S1**), our prediction efforts focused solely on response to the six leading drug classes. All cohorts underwent consistent pre-processing and normalization to ensure data uniformity. Expression data from different platforms were rank-normalized before downstream analyses (**Methods**).

To characterize transcriptomic variability among cohorts treated receiving the same drug class, we performed principal components analysis (PCA) on the rank-normalized pre-treatment gene expression data. This analysis revealed substantial variability across cohorts receiving the similar treatment with the same drug class (**Supplementary Figure S1**). To quantify the degree of clinical heterogeneity for each drug class, we employed a weighted index ($H_{w,d}$) that integrates cancer-type diversity, cohort spread, and treatment composition (**Methods**, **Supplementary Table S2**). Among all treatment groups, the bevacizumab cohorts exhibited the highest heterogeneity ($H_{w,d} = 0.81$), whereas ICB-melanoma cohorts showed the lowest heterogeneity ($H_{w,d} = 0.07$).

Using the uniformly processed rank-normalized expression data, we first tested whether expression levels of known drug targets alone could predict treatment response, without any model training. We then developed EXPRESSO, a drug-specific supervised LASSO (L1-regularized) logistic regression framework that leverages the full transcriptome while incorporating biological priors. EXPRESSO integrates two variants of biological information as input features in its LASSO model: EXPRESSO-T, where only the drug's known target gene is included as a penalty-free feature, and EXPRESSO-B, which augments this setup by also including context-specific biomarkers identified through a nested discovery procedure (**Supplementary Materials S2**). Known drug targets included *CD274* (anti–PD-1 inhibitors; separate models for ICB-melanoma and ICB-non-melanoma), *ERBB2* (trastuzumab), *VEGFA* (bevacizumab), *BRAF* (BRAF inhibitors), *TUBB1* (paclitaxel), and *TOP2A* and *TYMS* (FAC/FEC chemotherapy) (**Figure 1B**). As controls, we built two baseline models: (i) a 'vanilla' supervised LASSO model without drug target(s) or biomarker priors, to assess the added value of biological information, and (ii) an unsupervised target-only model. To mimic real-world applications and minimize overfitting, we used a Leave-One-Cohort-Out Cross-Validation (LOCO-CV) framework. Model performance was assessed using ROC-AUC and odds ratios (**Methods**).



B

Drug	Target(s)	Number of Cohorts	Total Sample Number	Number of Responder	Number of Non-responder
ICB-melanoma	CD274	9	336	143	193
ICB-non-melanoma	CD274	14	607	211	396
Trastuzumab	ERBB2	13	494	228	266
Bevacizumab	VEGFA	8	377	220	157
BRAFi	BRAF	7	192	107	85
Paclitaxel	TUBB1	9	933	290	643
Chemo-FAC-FEC	TOP2A, TYMS	9	790	240	550

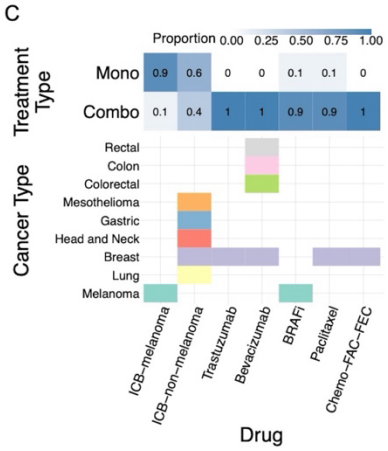


Figure 1: ***Study Overview and EXPRESSO workflow. (A) The workflow illustrating the evaluation of pre-treatment transcriptomic data for predicting patient response to cancer therapy.*** *The analysis consists of two main components - first, data processing and curation of large number of clinical trial cohorts, and second, the development of EXPRESSO – an AI-based, biologically grounded, drug-specific predictor of therapy response from pre-treatment tumor transcriptomes. EXPRESSO integrates two variants of biological information as input features in its LASSO model: EXPRESSO-T, which uses only the drug's known target gene as a penalty-free feature, and EXPRESSO-B, which augments this setup by also including context-specific biomarkers. Based on our study findings, EXPRESSO-T is recommended for*

combination therapy settings, whereas EXPRESSO-B performs optimally for monotherapy. ***(B) Summary table of the clinical trial cohorts used in this study****, and* ***(C) Characteristics of different drug classes in terms of treatment type*** *(either mono or combination therapy, the darker blue corresponds to higher fraction of cohorts treated with combination therapy and the lighter blue corresponds to the fraction of cohorts with monotherapy) and cancer types of the patients in each drug class. Color codes for different cancer types are: melanoma – sea green, lung – yellow, breast – lavender, head and neck – red, gastric – steel blue, mesothelioma – orange, colorectal – light green, colon – pink, and rectal – gray.*

Target expression alone demonstrates meaningful predictive power only in select monoclonal antibody, immunotherapy, and targeted therapies

We assessed whether expression of the drug target gene(s) alone (unsupervised mode) could predict treatment response across the seven sets of treatment-type cohorts (**Figure 2A-G**). Patients within each cohort were ranked based on the rank-normalized expression of each drug target gene, under the hypothesis that its expression might correlate with response. Across all drugs, the predictive performance of target gene expression alone varied widely, with mean ROC-AUCs ranging from 0.47 to 0.69. These findings suggest that while target gene expression may carry some predictive signal, it is insufficient for robust prediction.

Notably, monoclonal antibody therapies—whether targeted or immunotherapeutic—showed a meaningful association between target gene expression and clinical response. For anti–PD-1 immune checkpoint blockade (ICB) therapies, *CD274* (encodes PD-L1), expression showed moderate predictive power across 9 melanoma cohorts (mean ROC-AUC 0.65, empirical p-value = 0.0023), with *PDCD1* (PD-1) expression performing similarly (mean ROC-AUC 0.64; **Figure 2A**). In 14 non-melanoma ICB cohorts—including lung, gastric, breast, head and neck, and mesothelioma—*CD274* expression again predicted response modestly well (mean ROC-AUC 0.69, empirical p-value = 0.0004; **Figure 2B, 2H**), slightly outperforming *PDCD1* (mean ROC-AUC 0.66), consistent with prior reports[74–76]. For trastuzumab-treated breast cancer patients, higher *ERBB2* expression predicted therapeutic response, achieving a mean ROC-AUC of 0.68 (empirical p-value = 0.0002) across 13 cohorts (**Figure 2C**). Similarly, *VEGFA* expression modestly predicted response to bevacizumab in 8 cohorts across triple-negative breast,

colorectal, rectal, and colon cancers, also yielding a mean ROC-AUC of 0.68 (empirical p-value = 0.0002; **Figure 2D**).

In contrast, *BRAF* expression did not predict response to BRAF inhibitors in 7 melanoma cohorts (mean ROC-AUC 0.49, empirical p-value = 0.49; **Figure 2E**). This is not surprising because *BRAF* inhibitors target an aberrant protein with a p.V600 mutation[77]. Likewise, *TUBB1*, the known target of paclitaxel[78,79], showed no predictive power across 9 breast cancer cohorts (mean ROC-AUC 0.47, empirical p-value = 0.73; **Figure 2F**). For FAC-FEC chemotherapy, the expression of *TOP2A* (primary target of anthracyclines, like doxorubicin and epirubicin)[80] and *TYMS* (primary target of 5-fluorouracil)[81] showed moderate predictive performance (mean ROC-AUCs of 0.61 separately for each single gene across 9 cohorts, empirical p-value = 0.011 for *TOP2A* and empirical p-value = 0.0086 for *TYMS*; **Figure 2G**). Empirical p-values are based on control analyses using 10,000 randomly selected genes (**Methods**).

Together, these findings reveal that monoclonal antibody efficacy is more tightly linked to target gene expression, possibly due to the super-linear dependence[82] of antibody binding on target abundance. However, target expression alone is generally insufficient for accurate patient stratification. To go beyond target gene expression, we next opted for whole transcriptome analysis rather than limited gene panels, selecting genes measured in at least three cohorts for any specific treatment type to ensure comparability across datasets.

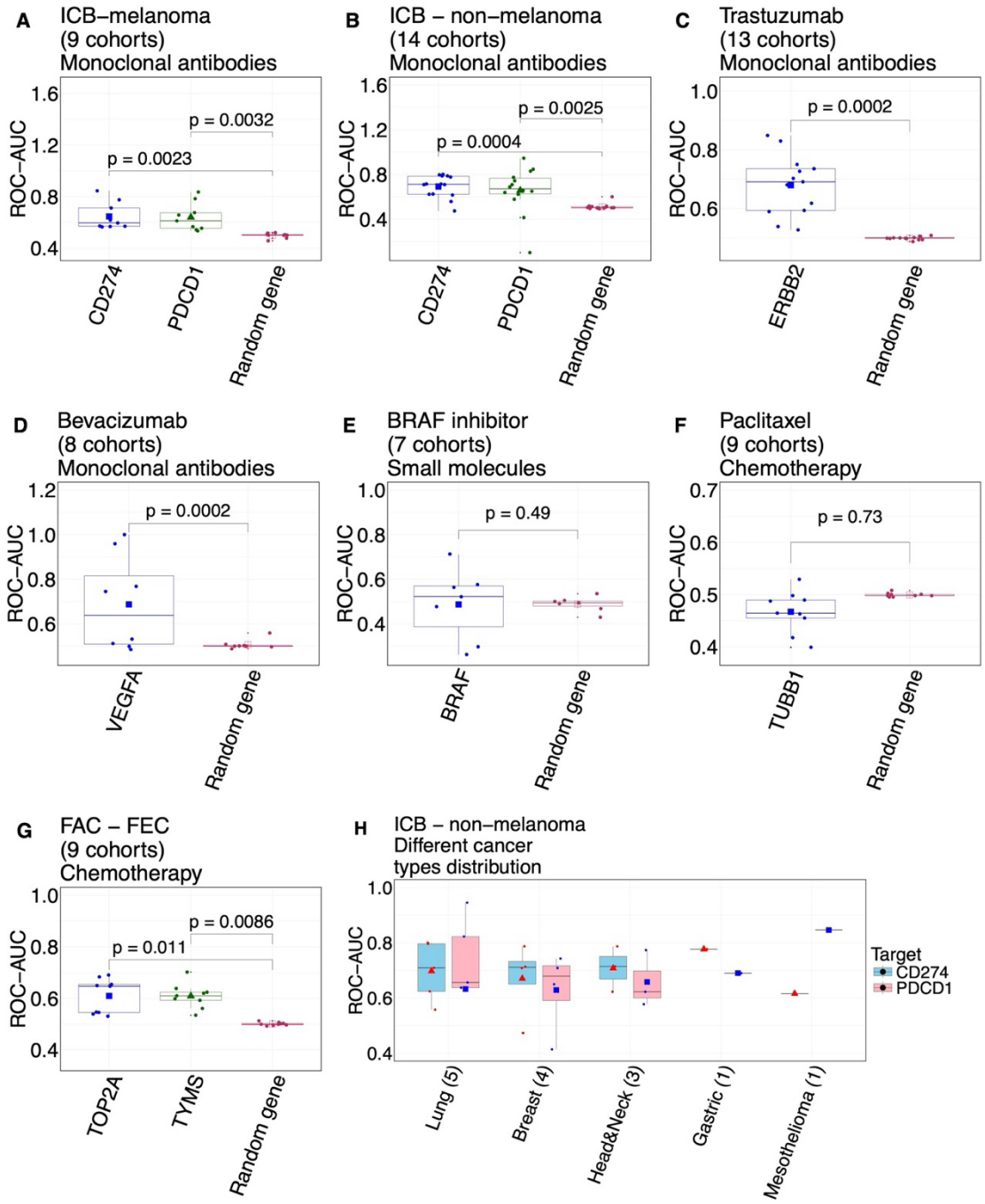


Figure 2: ***Drug response prediction based on target gene expression in an unsupervised approach.*** *(A-G) Box plots showing ROC-AUC scores of the expression of target genes and median ROC-AUCs of corresponding random controls for all six drug classes. For immunotherapy, we separated melanoma and non-melanoma cohorts. Empirical p-values comparing target expression*

with random controls (10,000 randomly selected genes) are shown. (H) Box plots showing ROC-AUC scores of the expression of target genes for ICB-non-melanoma for different cancer types.

EXPRESSO effectively predicts patients' treatment responses for seven different drug regimens

We developed seven drug-specific, biologically informed, EXPRESSO LASSO models using whole-transcriptome features with regularization. Biologically informed priors were incorporated either as target gene(s) alone (EXPRESSO-T) or combined with differentially expressed (DE) biomarkers from the training data (EXPRESSO-B, **Methods**), both treated as penalty-free features. Model performance was evaluated using a LOCO-CV strategy.

Overall, the two EXPRESSO variants achieved comparable performance (**Supplementary Figure S2**), except in ICB-melanoma, where EXPRESSO-B significantly outperformed EXPRESSO-T ($p = 0.02$, one-sided DeLong's test). Notably, ICB-melanoma exhibited the lowest clinical heterogeneity index ($H_{w,d} = 0.07$) among all drugs (**Supplementary Table S2**). In this setting, EXPRESSO-B identified additional biologically relevant genes - such as *IRF1*[83,84] and *UBD*[85,86] - beyond EXPRESSO-T features, thereby enhancing prediction accuracy (**Supplementary Materials S1, Supplementary Figure S3**). For downstream analyses, we retained only the variant with the higher mean ROC-AUC between the two modes.

Across all seven drug-specific models, EXPRESSO achieved mean ROC-AUC values ranging from 0.64 for BRAF inhibitors to 0.73 for trastuzumab (**Figures 3A-G**) and median odds ratios from 2.4 for bevacizumab to 4.6 for ICB-melanoma (**Figure 3H**). Compared to two baselines—a supervised vanilla model without biological priors and an unsupervised target-only model - the biologically informed EXPRESSO models consistently outperformed both (**Figure 3**). For bevacizumab, although the mean ROC-AUCs for EXPRESSO-T and the unsupervised target only model were identical (0.69), EXPRESSO-T exhibited markedly lower inter-cohort variability and significantly higher paired within-cohort ROC-AUCs (one-sided DeLong's p-value = 0.03; **Figure 3D**), reflecting improved consistency across datasets. Performance gains were most pronounced for targeted and immunotherapies, whereas chemotherapy models showed minimal improvement, reflecting limited predictive signal from target genes.

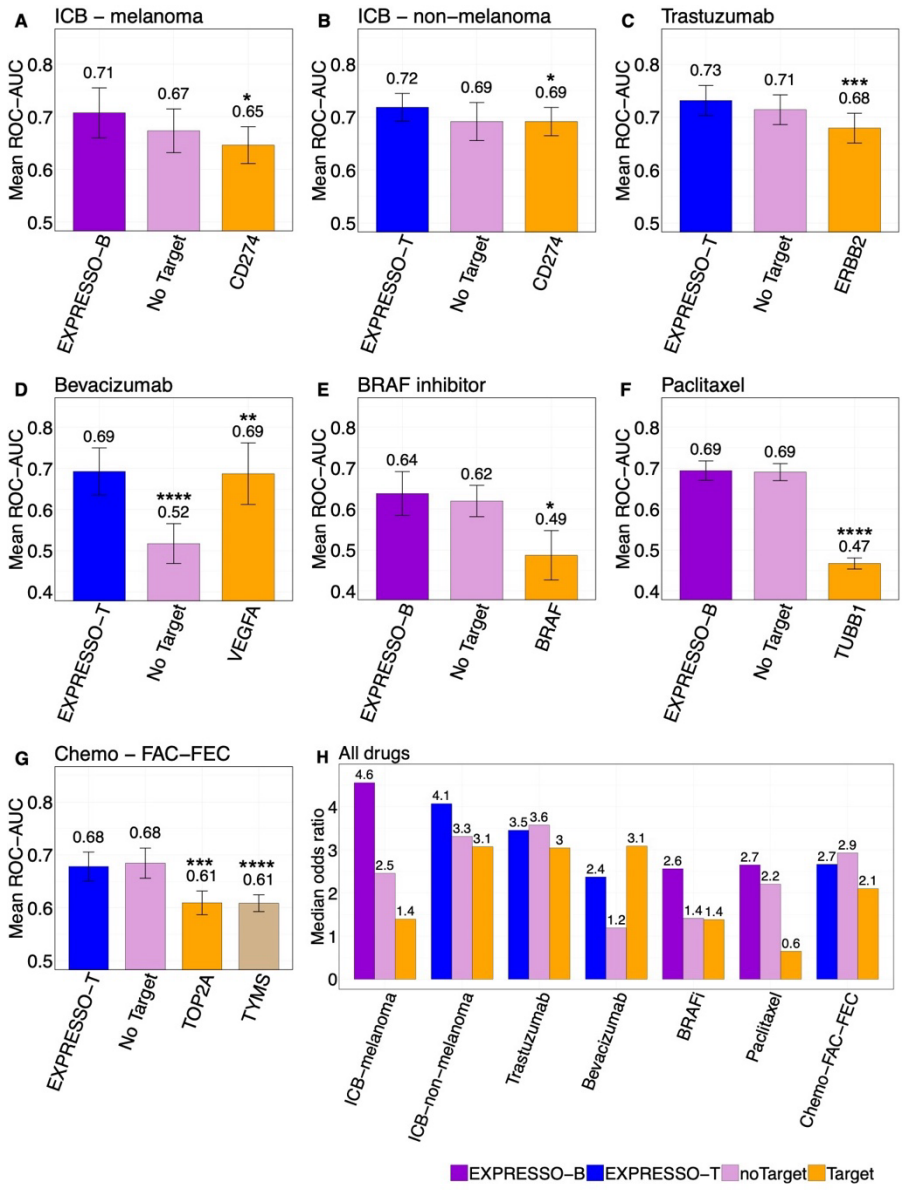


Figure 3: ***Drug response prediction of EXPRESSO compared with the two comparator baseline models.*** *(A–G) Bar plots of mean ROC-AUC with standard error bars for all seven drug-specific EXPRESSO models (EXPRESSO-T in blue and EXPRESSO-B in dark violet), vanilla supervised models without biological priors (No Target, plum), and unsupervised target-only predictors (orange for the first target and tan for the second target). The best-performing EXPRESSO variant (by higher mean ROC-AUC between EXPRESSO-T and EXPRESSO-B) is shown for each drug, as the two EXPRESSO variants achieved comparable performances, except ICB-melanoma (Supplementary Figure S2). Significant p-values (one-sided DeLong's test*

*across all cohorts for a drug) between EXPRESSO and no target supervised model or between EXPRESSO and unsupervised target model are denoted by an asterix (*) with significance levels as follows: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$, **** $p \leq 0.001$. (H) Median odds ratios for the seven drug-specific EXPRESSO models are shown in the bar plot, using the same color scheme as the preceding figure. For Chemo-FAC-FEC, the reported value for the target represents the median across all cohorts predicted by the two target genes.*

EXPRESSO benchmarks favorably against 20 published gene signatures and transcriptomic predictors

We benchmarked EXPRESSO against a comprehensive set of previously published gene signatures and transcriptomic predictors for treatment response across immune checkpoint blockade (ICB), chemotherapy, and breast cancer therapies. These included ICB-specific signatures such as ADO[87], Myeloid DC[88], TGFB_Mariathasan[89], APM_Thompson[90], APM_Wang[91], B_cell_Budczies and Macrophage[92], CD8_SF[93], CYT[94], Immune Cells[95], IMPRES[96], IPRES[97], COX_IS[98], and TIDE[99]; proliferation or oncogenic pathway-related markers such as KDM5A[100] and MKI67[101]; and breast cancer-specific predictors including MammaPrint[102,103], OncotypeDX[104], Immuno-Oncology signature[105], and an 11-gene signature[106].

Many of these signatures were originally developed in specific therapeutic or disease contexts (e.g., ICB response, chemotherapy sensitivity, or breast cancer prognosis) and were not designed for every drug evaluated here. Nonetheless, we applied the same set of 20 signatures across all six drug classes to establish a uniform benchmarking framework. Across drug-specific models, the best-performing EXPRESSO variant (the mode with the higher mean ROC-AUC) consistently outperformed these existing gene signatures in terms of mean ROC-AUC (**Figure 4**). In ICB-melanoma, the improvement was driven by EXPRESSO-B, whereas EXPRESSO-T did not exceed all the comparators (**Supplementary Figure S4**). For other drugs, both EXPRESSO variants outperformed published signatures. These results highlight the strength of target-aware supervised modeling with stringent biomarker selection and underscore the importance of treatment-specific predictors in precision oncology.

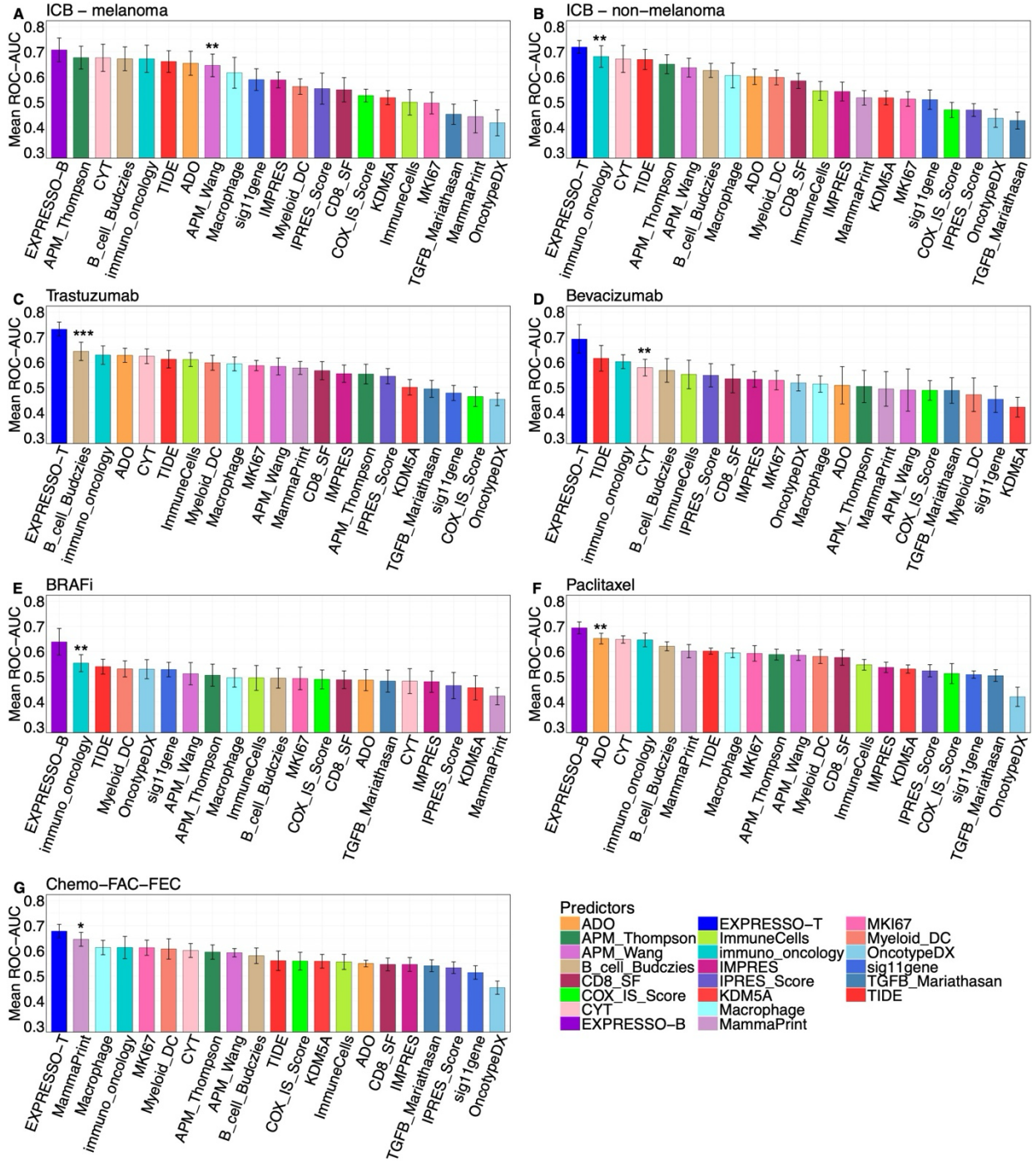


Figure 4: ***Comparative performance of best-performing EXPRESSO variant for each drug class against existing gene signatures and methods.*** *Bar plots show the mean ROC-AUC values*

*for each predictor across drug-specific models, ordered from highest to lowest mean performance. Significant p-values (one-sided DeLong's test across all cohorts for a drug) between EXPRESSO variant (by higher mean ROC-AUC between EXPRESSO-T and EXPRESSO-B) and the first significantly different predictor denoted by Asterix (*) with significant levels as follows: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; subsequent predictors with lower performance are also significantly different. While many of the published signatures were originally developed for specific therapeutic or disease contexts, we applied the same set of 20 signatures uniformly across all drugs to enable consistent benchmarking. The best-performing EXPRESSO variant is shown for each drug, as the two EXPRESSO variants achieved comparable performances, except ICB-melanoma (Supplementary Figure S2).*

Supervised prediction capability limit of transcriptomic data varies across different treatment types

To estimate the upper bound of supervised predictive performance and assess the impact of training data size, we conducted incremental training experiments (**Figure 5**). For each drug, we sequentially added training cohorts one at a time (excluding the test cohort) in various orders and retrained the best performing EXPRESSO variant (the mode with the higher mean ROC-AUC) at each step of each order (**Methods**). Across drug types, we observed consistent increases in ROC-AUC as additional cohorts were incorporated, although in some drug types, the improvements diminished after a certain point—suggesting saturation of predictive performance for some drugs.

In the analyses shown in **Figure 5**, each training subset with the same number *t* of training cohorts was treated equivalently, but these subsets may have rather different sample sizes. Therefore, we also made scatter plots of sample size (x-axis) against ROC-AUC (y-axis) for fixed numbers of training cohorts (**Methods, Supplementary Figure S5**). Reassuringly, ROC-AUC is consistently positively correlated with sample size, significantly for some drugs and insignificantly for other drugs. The lines of best fit have small coefficients for x suggesting that where EXPRESSO prediction accuracy has not plateaued, the addition of new cohorts of roughly average size (relative to the current cohorts) would help.

In summary, these experiments suggest that predictive performance may have reached its upper bound for trastuzumab, ICB-melanoma, and ICB-non-melanoma where we have sufficient (at

least nine) cohorts. In contrast, models for bevacizumab, BRAF inhibitors, and for chemotherapy cohorts have not yet plateaued.

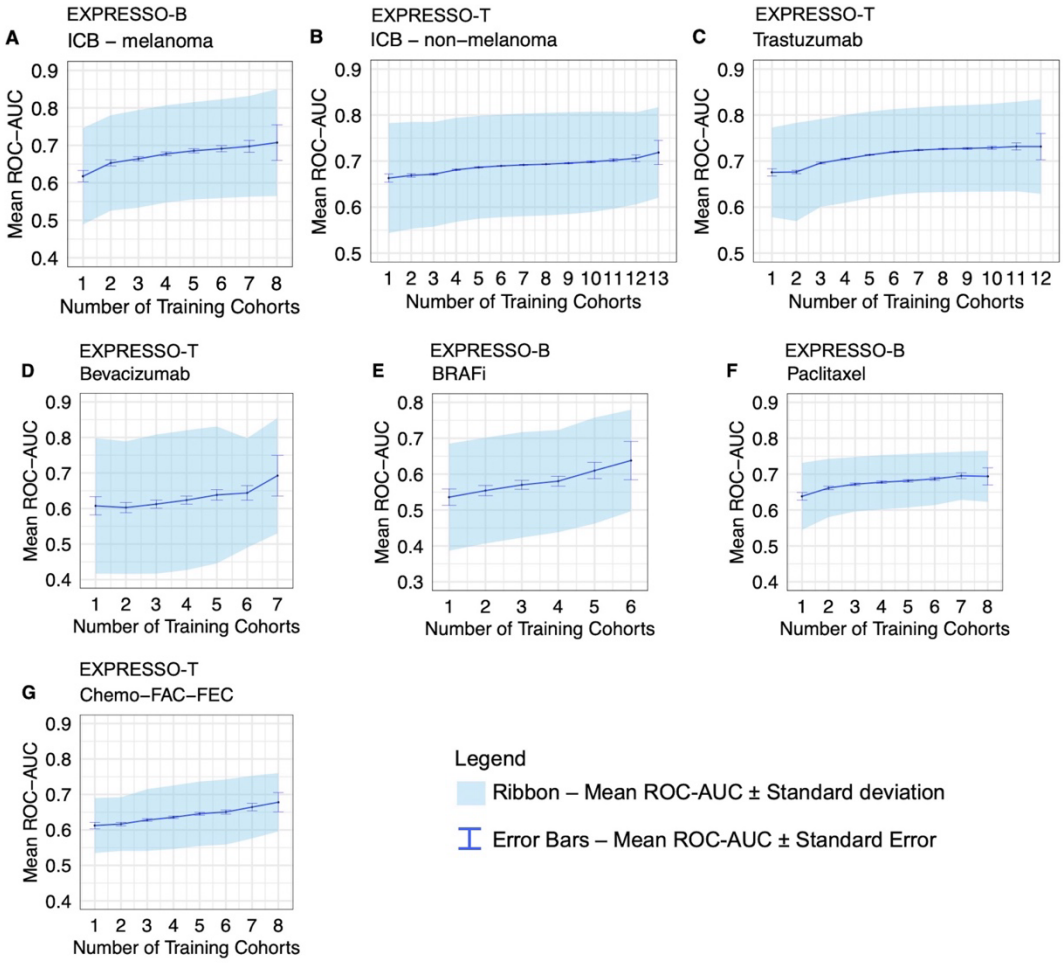


Figure 5: *Line plot with standard deviation as ribbon (sky blue) and mean standard error as error bars for the upper bound of transcriptome-based prediction for seven drug-specific best-performing EXPRESSO variants (either EXPRESSO-T or EXPRRESSO-B, as noted on top of the plots) for each drug class.*

DISCUSSION

In this study, we evaluated the predictive potential of bulk transcriptomic data across diverse cancer therapies using supervised machine learning frameworks. By compiling and uniformly

processing 69 clinical cohorts (3,729 patients) covering six frontline drug classes - anti-PD-1 immunotherapy (melanoma and non-melanoma), trastuzumab, bevacizumab, BRAF inhibitors, paclitaxel, and FAC/FEC chemotherapy - we developed an approach to assess transcriptomics in a pan-cancer setting. Notably, all datasets are publicly available (but some require dbGaP permission) and the Expresso code is available on zenodo.

We showed that expression of drug targets can predict treatment response quite reasonable (mean ROC-AUC of 0.68-0.69) for selected monoclonal antibody therapies (trastuzumab, bevacizumab, and anti-PD-1 inhibitors for non-melanoma cancers), but not for small-molecule BRAF inhibitors, paclitaxel, or FAC/FEC chemotherapy. In melanoma, PD-L1 expression was not a definitive predictor of response[107–109], though we observed modest predictive value (mean ROC-AUC 0.65). These observations suggest that monoclonal antibody efficacy may depend more directly on target abundance[82], whereas chemotherapy and small-molecule inhibitors often act through mutated proteins or multiple/off-target mechanisms, making single-gene transcript levels an imperfect proxy for drug activity[110]. While previous studies have reported target-specific associations in narrower drug–cancer-cohort contexts[40,44,111–113], our work provides a broader evaluation across multiple cancer types, therapies, and clinical cohorts.

We developed EXPRESSO, a drug-specific LASSO framework, integrating biological priors with whole transcriptome features to capture predictive signals beyond target expression while reducing risk of overfitting using LOCO-CV strategy. It incorporates biological information in two variants: EXPRESSO-T, which includes only the drug's target gene(s), and EXPRESSO-B, which additionally includes context-specific biomarkers, as penalty-free features for both. Both variants performed comparably for all the drugs, except in ICB-melanoma, where EXPRESSO-B outperformed EXPRESSO-T ($p=0.02$, one-sided DeLong's test). While achieving higher performance than 20 published gene signatures and prediction tools, overall, EXPRESSO achieved moderate response prediction across several drug types. In many cases, it consistently outperformed the baseline supervised vanilla model without biological priors, demonstrating the value of incorporating biological knowledge.

This study has several limitations. First, using both RNA-seq and microarray datasets may introduce technical variability, though batch correction did not improve performance. Second,

anti-PD-1 and anti-PD-L1 therapies were analyzed together to increase cohort sizes, despite their known efficacy differences[114–116]. Third, importantly, many therapies are often delivered in combinations, but we modeled only the primary treatment to maximize cohort inclusion. Finally, we posed prediction dichotomously ("Will patient *p* respond to treatment *t*?"), whereas tumor boards face a combinatorial decision across multiple options.

Perhaps the most important finding of this study is a negative one: at least in some treatment types, prediction accuracy plateaued and did not improve with additional training cohorts. This suggests that supervised bulk transcriptomics alone may be insufficient for robustly predicting patients' response in highly heterogeneous cohorts, especially for the case of non-targeted therapies. Our findings highlight the inherent limitations of supervised brute-force learning for certain treatments, while underscoring the potential of incorporating additional data and deeper mechanistic modeling of tumor heterogeneity to improve transcriptomics-based predictors. Future directions are of two conceptual dimensions: (a) developing bulk expression models for still larger and more clinically homogenous datasets, that, importantly, better model and incorporate the underlying biological mechanisms of treatment resistance and response, and (b) bringing additional omics data to the mix, include integrating DNA mutation data, prior treatments information, clinical metadata, and richer multi-omics (e.g., single-cell transcriptomics, proteomics, spatial immune profiling and others), to provide a more comprehensive view of therapeutic response mechanisms. Increasing access to large, well-annotated, and diverse clinical trial cohorts, with such deeper molecular profiling, will hence be important for advancing data-driven precision oncology. Fortunately, recent developments in AI-based inference of such types of data in a fast and cost-effective manner directly from the tumor pathology slides (e.g., DeepPT[117], Path2Omics[118], Path2Space[119]) may facilitate the latter in a readily accessible and cost effective manner.

MATERIALS and METHODS

Data collection and processing

We performed three systematic searches for suitable articles in multiple medical research databases: Embase [Nov, 2023], PubMed [Nov, 2023], and Trialtrove [Aug 5, 2023 data freeze,

total 93,161 oncology clinical trials][120] databases. For the Embase and Trialtrove searches, we required that the study must be a (prospective) clinical trial, but we did not impose the clinical trial requirement in our PubMed search (**Supplementary Table S3**). Our search criteria were designed to find patients' pre-treatment bulk transcriptomic (RNAseq or microarray) data in studies where i) patients received targeted therapy (either monoclonal antibody or small molecule), immunotherapy, or only chemotherapy for a drug of interest and ii) the response data after the treatment are available (**Supplementary Table S1**). We filtered out cohorts with fewer than ten samples and required that each cohort partition the patients into two annotated two groups - responders and non-responders. Clinical trials with multiple arms (different treatment) were treated as separate cohorts. To identify the patient response labels (either via RECIST [Response Evaluation Criteria in Solid Tumors] or pCR [pathological Complete Response] labels), we followed the same criteria for responder/non-responder as mentioned in the original publications or in the metadata from dbGaP or GEO for each cohort.

For RNAseq data from GEO[121] (https://www.ncbi.nlm.nih.gov/geo/), we downloaded the RNAseq normalized counts TPM data directly. For microarray data from GEO, we also downloaded the data directly from GEO. While selecting genes across datasets, we required the gene symbols to be the same – we might miss some alias genes due to gene mapping issues. We used the collapseRows function from the WGCNA package[122,123] with the MaxMean method (probe with maximum value of mean expression across all samples) to resolve the duplicated genes with multiple probes in the microarray.

To analyze gene expression data from dbGAP[124], we first downloaded raw FASTQ files. Quality control (QC) of the raw sequencing reads was performed using FastQC to assess sequencing quality and detect potential issues. Low-quality reads and bases were filtered out using standard QC parameters such as read quality scores (Q30). The remaining high-quality reads were then processed using Salmon[125] with its default settings to estimate transcript-level TPM (Transcripts Per Million) and read counts. Subsequently, the tximport R package[126], with its default settings, was used to summarize transcript-level abundances from Salmon's output into gene-level TPM and count matrices.

We used rank normalized (by patient across genes) expression data from each cohort for further downstream analyses.

Clinical heterogeneity index calculation

To compare the degree of heterogeneity across drugs, we computed a clinical heterogeneity index, $H_{w,d}$ that integrates cancer-typer diversity, cohort spread and treatment mix for a given drug.

Cancer-type diversity: $\boldsymbol{H_{cancer,d}}$

$$H_{cancer,d} = \frac{C_d - 1}{C_{max} - 1}$$

Where, $C_d$is the number of distinct cancer types tested for drug d, and C_{max} is the maximum number of types of cancer for any drug in the list. This normalizes cancer-type spread between 0 (drug with single cancer type) and 1 (drug with most diverse cohorts).

Cancer-cohort balance index: $\boldsymbol{H_{CC,d}}$

$$H_{CC,d} = H_{cancer,d} \times (1 - e^{-\alpha.Cohort_d})$$

Where, $(1 - e^{-\alpha.Cohort_d})$ is a scaling function, where adding cohorts increases "balance" but with diminishing returns. Alpha (α) is the saturation rate parameter that controls how quickly the cohort correction approaches 1.

Treatment mix: $\boldsymbol{H_{treatment,d}}$

$$H_{treatment,d} = \frac{T_{combo,d}}{T_d}$$

where, $T_{combo,d}$ is the number of cohorts with combination treatment for a drug d and T_d is the total number of cohorts for that drug d. Values near 1 indicate predominance of combination treatment (greater heterogeneity), while values near 0 indicate mostly monotherapy (greater homogeneity).

The final Heterogeneity index is defined in a weighted manner as:

$$H_{w,d} = w_1.H_{CC,d} + w_2.H_{treatment,d}$$

where the contribution of cancer-cohort balance for a drug d, $w_1 = 0.4$ and the contribution for treatment mix, $w_2 = 0.6$. We assigned greater weight (0.6) to the treatment mix component based on our intuition that variability among overall treatment regimen is more important than the cancer-cohort balance index.

Generating LASSO Logistic Regression Models

We utilized the glmnet R package[127] to create LASSO logistic regression models from the rank-normalized features (details in **Supplementary Materials S2**).

Leave-One-Cohort-Out Cross-Validation (LOCO-CV) Framework

To ensure robust and generalizable evaluation, we employed a Leave-One-Cohort-Out Cross-Validation (LOCO-CV) framework for the supervised learning (details in **Supplementary Materials S2**).

EXPRESSO

EXPRESSO is a drug-specific AI-based supervised learning framework specifically tailored for transcriptome-based drug response prediction. EXPRESSO leverages LASSO regression with L1 regularization, chosen for its capability to perform automatic feature selection and mitigate overfitting. EXPRESSO incorporates the entire transcriptome as input for feature selection and model training. It further integrates biological knowledge in the model using two strategies: (a) keeping the target gene(s) only as penalty-free (EXPRESSO-T, **Supplementary Materials S2**), or (b) using additional context-specific genes together with the target gene(s) as penalty-free (EXPRESSO-B, **Supplementary Materials S2**). Biomarkers were identified in a drug-specific manner using nested LOCO-CV framework (**Supplementary Materials S2**). For downstream analyses, we retained only the variant with the higher mean ROC-AUC between the two modes. We used two baseline models – a supervised 'vanilla' model without any biological priors and an unsupervised target only model to compare the performance of EXPRESSO (**Supplementary Materials S2**).

Performance measures

Model performance in each fold of the LOCO-CV framework was evaluated using both discrimination and effect size metrics. The primary metric was the area under the ROC curve (ROC-AUC), calculated from predicted probabilities on the held-out cohort. We also computed odds ratios (OR) by comparing response rates between patients predicted positive (score ≥ 0.5) and the remainder of the cohort, with 95% confidence intervals from Fisher's exact test. For each drug, mean ROC-AUC and median OR across folds were reported, with drug-class–level summaries based on median OR to reduce outlier effects. Statistical comparisons between models used paired Wilcoxon signed-rank tests and DeLong's test, where applicable (details in **Supplementary Materials S2**).

Incremental Training for Upper-Bound Performance Estimation

To estimate the upper bound of predictive accuracy and assess the effect of training cohort size, we performed incremental training within the LOCO-CV framework. For each test cohort, models were retrained iteratively by adding one training cohort at a time (from a single cohort up to all available cohorts, excluding the test cohort) and evaluated on the fixed test cohort. This stepwise accumulation allowed us to track how ROC-AUC changed with increasing training data and to approximate performance saturation points. To further evaluate scaling, we also plotted ROC-AUC against the number of training patients (rather than cohorts) across runs (**Supplementary Fig. S5**) (details in **Supplementary Materials S2**).

DATA/CODE AVAILABILITY

All datasets are publicly available, but some require dbGaP permission (**Supplementary Table S1**). The code is available on zenodo (https://doi.org/10.5281/zenodo.17422273).

ACKNOWLEDGEMENTS

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH), National Cancer Institute. The contributions of the NIH author(s) were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered Works of the United States Government. However, the findings

and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services. This work used the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). We thank Stacy Brody from the NIH Library and Gal Dinstag from Pangea, for their help with this work. The authors used ChatGPT 4o and 5 to improve the quality of writing of this manuscript.

DECLARATION OF POTENTIAL CONFLICTS OF INTEREST

E.R. is a cofounder of MedAware Ltd. and a cofounder (divested) and nonpaid scientific consultant of Pangea Therapeutics. E.R. is a member of the scientific advisory board of GSK Oncology, the WIN consortium and the ProCan program. All other authors have declared no conflicts of interest.

AUTHOR CONTRIBUTIONS

L.R.P.: Data collection, data pre-processing, data curation, conceptualization, software, methodology, formal analysis, visualization, writing the draft. **E.M.G.:** Software, methodology, formal analysis, visualization. **N.U.N:** Conceptualization, methodology, writing the draft. **S.M.:** Data pre-processing, formal analysis. **S.P.:** Data collection, Formal analysis. **T.C.:** Data pre-processing, software. **E.M.C.:** Data pre-processing. **T-G.C:** Data pre-processing. **S.R.D.:** Data collection. **Y. K.:** Data collection. **E.D.S.:** Methodology. **P.S.R.:** Data collection. **D-T.H.:** Methodology review. **A.A.S.:** Data collection, data acquisition, supervision, edited the original draft and carried out the revisions. **E.R.:** Conceptualization, data collection, supervision, project administration, funding acquisition, methodology, edited the original draft and carried out the revisions. All authors critically reviewed and approved the final version of the manuscript.

REFERENCES

1. Jameson, G. S. *et al.* Response rate following albumin-bound paclitaxel plus gemcitabine plus cisplatin treatment among patients with advanced pancreatic cancer: a phase 1b/2 pilot clinical trial. *JAMA Oncol.* **6**, 125–132 (2020).
2. Hyman, D. M., Taylor, B. S. & Baselga, J. Implementing genome-driven oncology. *Cell* **168**, 584–599 (2017).
3. Van Allen, E. M. *et al.* Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nat. Med.* **20**, 682–688 (2014).
4. PCAWG Transcriptome Core Group *et al.* Genomic basis for RNA alterations in cancer. *Nature* **578**, 129–136 (2020).
5. Liu, B., Zhou, H., Tan, L., Siu, K. T. H. & Guan, X.-Y. Exploring treatment options in cancer: tumor treatment strategies. *Signal Transduct. Target. Ther.* **9**, 175 (2024).
6. Min, H.-Y. & Lee, H.-Y. Molecular targeted therapy for anticancer treatment. *Exp. Mol. Med.* **54**, 1670–1694 (2022).
7. Zamani, M. R. & Šácha, P. Immune checkpoint inhibitors in cancer therapy: what lies beyond monoclonal antibodies? *Med. Oncol.* **42**, 273 (2025).
8. Zhong, L. *et al.* Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. *Signal Transduct. Target. Ther.* **6**, 201 (2021).
9. He, X. & Xu, C. Immune checkpoint signaling and cancer immunotherapy. *Cell Res.* **30**, 660–669 (2020).

10. Shiravand, Y. *et al*. Immune checkpoint inhibitors in cancer therapy. *Curr. Oncol*. **29**, 3044–3060 (2022).

11. Justiz-Vaillant, A., Pandit, B. R., Unakal, C., Vuma, S. & Akpaka, P. E. A comprehensive review about the use of monoclonal antibodies in cancer therapy. *Antibodies Basel Switz*. **14**, 35 (2025).

12. Tsao, L.-C., Force, J. & Hartman, Z. C. Mechanisms of therapeutic antitumor monoclonal antibodies. *Cancer Res*. **81**, 4641–4651 (2021).

13. Zahavi, D. & Weiner, L. Monoclonal antibodies in cancer therapy. *Antibodies* **9**, 34 (2020).

14. Liu, G.-H., Chen, T., Zhang, X., Ma, X.-L. & Shi, H.-S. Small molecule inhibitors targeting the cancers. *MedComm* **3**, e181 (2022).

15. Yesilkanal, A. E., Johnson, G. L., Ramos, A. F. & Rosner, M. R. New strategies for targeting kinase networks in cancer. *J. Biol. Chem*. **297**, 101128 (2021).

16. Anand, U. *et al*. Cancer chemotherapy and beyond: current status, drug candidates, associated risks and progress in targeted therapeutics. *Genes Dis*. **10**, 1367–1401 (2023).

17. Nesic, K., Parker, P., Swisher, E. M. & Krais, J. J. DNA repair and the contribution to chemotherapy resistance. *Genome Med*. **17**, 62 (2025).

18. van den Boogaard, W. M. C., Komninos, D. S. J. & Vermeij, W. P. Chemotherapy side-effects: not all DNA damage is equal. *Cancers* **14**, 627 (2022).

19. Bedard, P. L., Hansen, A. R., Ratain, M. J. & Siu, L. L. Tumour heterogeneity in the clinic. *Nature* **501**, 355–364 (2013).

20. Dienstmann, R. *et al.* Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat. Rev. Cancer* **17**, 79–92 (2017).

21. Meric-Bernstam, F. *et al.* A decision support framework for genomically informed investigational cancer therapy. *J. Natl. Cancer Inst.* **107**, djv098 (2015).

22. Lin, V. T. G. & Yang, E. S. The pros and cons of incorporating transcriptomics in the age of precision oncology. *J. Natl. Cancer Inst.* **111**, 1016–1022 (2019).

23. Tsimberidou, A. M., Fountzilas, E., Bleris, L. & Kurzrock, R. Transcriptomics and solid tumors: the next frontier in precision cancer medicine. *Semin. Cancer Biol.* **84**, 50–59 (2022).

24. Zehir, A. *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **23**, 703–713 (2017).

25. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* **168**, 707–723 (2017).

26. Hoadley, K. A. *et al.* Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **158**, 929–944 (2014).

27. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

28. AlDoughaim, M. *et al.* Cancer biomarkers and precision oncology: a review of recent trends and innovations. *Clin. Med. Insights Oncol.* **18**, 11795549241298541 (2024).

29. Passaro, A. *et al.* Cancer biomarkers: emerging trends and clinical implications for personalized treatment. *Cell* **187**, 1617–1635 (2024).

30. Zhou, Y. *et al*. Tumor biomarkers for diagnosis, prognosis and targeted therapy. *Signal Transduct. Target. Ther.* **9**, 132 (2024).

31. Chakravarty, D. *et al*. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis. Oncol.* 1–16 (2017) doi:10.1200/PO.17.00011.

32. Suehnholz, S. P. *et al*. Quantifying the Expanding Landscape of Clinical Actionability for Patients with Cancer. *Cancer Discov.* **14**, 49–65 (2024).

33. Füzéry, A. K., Levin, J., Chan, M. M. & Chan, D. W. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin. Proteomics* **10**, 13 (2013).

34. Li, D. & Chan, D. W. Proteomic cancer biomarkers from discovery to approval: it's worth the effort. *Expert Rev. Proteomics* **11**, 135–136 (2014).

35. Chapman, P. B. *et al*. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N. Engl. J. Med.* **364**, 2507–2516 (2011).

36. Garon, E. B. *et al*. Pembrolizumab for the treatment of non-small-cell lung cancer. *N. Engl. J. Med.* **372**, 2018–2028 (2015).

37. Karapetis, C. S. *et al*. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N. Engl. J. Med.* **359**, 1757–1765 (2008).

38. Le, D. T. *et al*. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).

39. Mok, T. S. *et al*. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* **361**, 947–957 (2009).

40. Slamon, D. J. *et al*. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med*. **344**, 783–792 (2001).

41. Paik, S. *et al*. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med*. **351**, 2817–2826 (2004).

42. Sparano, J. A. *et al*. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N. Engl. J. Med*. **379**, 111–121 (2018).

43. Ayers, M. *et al*. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest*. **127**, 2930–2940 (2017).

44. Dinstag, G. *et al*. Clinically oriented prediction of patient response to targeted and immunotherapies from the tumor transcriptome. *Med N. Y. N* **4**, 15-30.e8 (2023).

45. Geeleher, P. *et al*. Discovering novel pharmacogenomic biomarkers by imputing drug response in cancer patients from large genomics studies. *Genome Res*. **27**, 1743–1751 (2017).

46. Lee, J. S. *et al*. Synthetic lethality-mediated precision oncology via the tumor transcriptome. *Cell* **184**, 2487-2502.e13 (2021).

47. Lee, J.-K. *et al*. Pharmacogenomic landscape of patient-derived tumor cells informs precision oncology therapy. *Nat. Genet*. **50**, 1399–1411 (2018).

48. Triantafyllidis, C. P. *et al*. A machine learning and directed network optimization approach to uncover TP53 regulatory patterns. *iScience* **26**, 108291 (2023).

49. Warren, A. *et al*. Global computational alignment of tumor and cell line transcriptional profiles. *Nat. Commun*. **12**, 22 (2021).

50. Geeleher, P., Cox, N. J. & Huang, R. S. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome Biol.* **15**, R47 (2014).

51. Bhinder, B., Gilvary, C., Madhukar, N. S. & Elemento, O. Artificial Intelligence in Cancer Research and Precision Medicine. *Cancer Discov.* **11**, 900–915 (2021).

52. Ma, J. *et al.* Few-shot learning creates predictive models of drug response that translate from high-throughput screens to individual patients. *Nat. Cancer* **2**, 233–244 (2021).

53. He, D., Liu, Q., Wu, Y. & Xie, L. A context-aware deconfounding autoencoder for robust prediction of personalized clinical drug response from cell-line compound screening. *Nat. Mach. Intell.* **4**, 879–892 (2022).

54. Kuenzi, B. M. *et al.* Predicting drug response and synergy using a deep learning model of human cancer cells. *Cancer Cell* **38**, 672-684.e6 (2020).

55. Partin, A. *et al.* Deep learning methods for drug response prediction in cancer: predominant and emerging trends. *Front. Med.* **10**, 1086097 (2023).

56. Shen, W. *et al.* Generalizable AI predicts immunotherapy outcomes across cancers and treatments. Preprint at https://doi.org/10.1101/2025.05.01.25326820 (2025).

57. Ling, A. & Huang, R. S. Computationally predicting clinical drug combination efficacy with cancer cell line screens and independent drug action. *Nat. Commun.* **11**, 5848 (2020).

58. Beaubier, N. *et al.* Integrated genomic profiling expands clinical options for patients with cancer. *Nat. Biotechnol.* **37**, 1351–1360 (2019).

59. Cuppen, E. *et al*. Implementation of whole-genome and transcriptome sequencing into clinical cancer care. *JCO Precis. Oncol.* **6**, e2200245 (2022).

60. Ianevski, A. *et al*. Single-cell transcriptomes identify patient-tailored therapies for selective co-inhibition of cancer clones. *Nat. Commun.* **15**, 8579 (2024).

61. Lazar, V. *et al*. A transcriptomics approach to expand therapeutic options and optimize clinical trials in oncology. *Ther. Adv. Med. Oncol.* **15**, 17588359231156382 (2023).

62. Pleasance, E. *et al*. Whole-genome and transcriptome analysis enhances precision cancer treatment options. *Ann. Oncol.* **33**, 939–949 (2022).

63. Rodon, J. *et al*. Genomic and transcriptomic profiling expands precision cancer medicine: the WINTHER trial. *Nat. Med.* **25**, 751–758 (2019).

64. Kato, S. *et al*. Real-world data from a molecular tumor board demonstrates improved outcomes with a precision N-of-one strategy. *Nat. Commun.* **11**, 4965 (2020).

65. Mundi, P. S. *et al*. A transcriptome-based precision oncology platform for patient-therapy alignment in a diverse set of treatment-resistant malignancies. *Cancer Discov.* **13**, 1386–1407 (2023).

66. Irmisch, A. *et al*. The tumor profiler study: integrated, multi-omic, functional tumor profiling for clinical decision support. *Cancer Cell* **39**, 288–293 (2021).

67. Borisov, N. & Buzdin, A. Transcriptomic harmonization as the way for suppressing cross-platform bias and batch effect. *Biomedicines* **10**, 2318 (2022).

68. Costello, J. C. *et al*. A community effort to assess and improve drug sensitivity prediction algorithms. *Nat. Biotechnol.* **32**, 1202–1212 (2014).

69. Guinney, J. *et al*. The consensus molecular subtypes of colorectal cancer. *Nat. Med*. **21**, 1350–1356 (2015).

70. Feng, J. *et al*. Metabolic gene signature for predicting breast cancer recurrence using transcriptome analysis. *Future Oncol*. **17**, 71–80 (2021).

71. Khatamian, A., Paull, E. O., Califano, A. & Yu, J. SJARACNe: a scalable software tool for gene network reverse engineering from big data. *Bioinforma. Oxf. Engl*. **35**, 2165–2166 (2019).

72. Wang, D. Y. *et al*. Fatal toxic effects associated with immune checkpoint inhibitors: a systematic review and meta-analysis. *JAMA Oncol*. **4**, 1721–1728 (2018).

73. Wu, Y. *et al*. Identification of the six-RNA-binding protein signature for prognosis prediction in bladder cancer. *Front. Genet*. **11**, 992 (2020).

74. De Sousa Linhares, A. *et al*. Therapeutic PD-L1 antibodies are more effective than PD-1 antibodies in blocking PD-1/PD-L1 signaling. *Sci. Rep*. **9**, 11472 (2019).

75. Taube, J. M. *et al*. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res*. **20**, 5064–5074 (2014).

76. Yang, F., Wang, J. F., Wang, Y., Liu, B. & Molina, J. R. Comparative analysis of predictive biomarkers for PD-1/PD-L1 inhibitors in cancers: developments and challenges. *Cancers* **14**, 109 (2021).

77. Castellani, G. *et al*. BRAF mutations in melanoma: biological aspects, therapeutic implications, and circulating biomarkers. *Cancers* **15**, 4026 (2023).

78. Jordan, M. A. & Wilson, L. Microtubules as a target for anticancer drugs. *Nat. Rev. Cancer* **4**, 253–265 (2004).

79. Kovács, P., Csaba, G., Pállinger, E. & Czaker, R. Effects of taxol treatment on the microtubular system and mitochondria of tetrahymena. *Cell Biol. Int.* **31**, 724–732 (2007).

80. Marinello, J., Delcuratolo, M. & Capranico, G. Anthracyclines as topoisomerase II poisons: from early studies to new perspectives. *Int. J. Mol. Sci.* **19**, 3480 (2018).

81. Kuremsky, J. G., Tepper, J. E. & McLeod, H. L. Biomarkers for response to neoadjuvant chemoradiation for rectal cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **74**, 673–688 (2009).

82. Dubacheva, G. V., Curk, T. & Richter, R. P. Determinants of superselectivity–practical concepts for application in biology and medicine. *Acc. Chem. Res.* **56**, 729–739 (2023).

83. Purbey, P. K. *et al.* Opposing tumor-cell-intrinsic and -extrinsic roles of the IRF1 transcription factor in antitumor immunity. *Cell Rep.* **43**, 114289 (2024).

84. Smithy, J. W. *et al.* Nuclear IRF-1 expression as a mechanism to assess 'capability' to express PD-L1 and response to PD-1 therapy in metastatic melanoma. *J. Immunother. Cancer* **5**, 25 (2017).

85. Hu, X. *et al.* Emerging role of ubiquitination in the regulation of PD-1/PD-L1 in cancer immunotherapy. *Mol. Ther. J. Am. Soc. Gene Ther.* **29**, 908–919 (2021).

86. Liu, J. *et al.* Targeting the ubiquitination/deubiquitination process to regulate immune checkpoint pathways. *Signal Transduct. Target. Ther.* **6**, 28 (2021).

87. Sidders, B. *et al.* Adenosine signaling is prognostic for cancer outcome and has predictive utility for immunotherapeutic response. *Clin. Cancer Res.* **26**, 2176–2187 (2020).

88. Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* **577**, 549–555 (2020).

89. Mariathasan, S. *et al.* TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).

90. Thompson, J. C. *et al.* Gene signature of antigen processing and presentation machinery predicts response to checkpoint blockade in non-small cell lung cancer (NSCLC) and melanoma. *J. Immunother. Cancer* **8**, e000974 (2020).

91. Wang, S., He, Z., Wang, X., Li, H. & Liu, X.-S. Antigen presentation and tumor immunogenicity in cancer immunotherapy response prediction. *eLife* **8**, e49020 (2019).

92. Budczies, J. *et al.* A gene expression signature associated with B cells predicts benefit from immune checkpoint blockade in lung adenocarcinoma. *Oncoimmunology* **10**, 1860586 (2021).

93. Sade-Feldman, M. *et al.* Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell* **175**, 998-1013.e20 (2018).

94. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).

95. Xiong, D., Wang, Y. & You, M. A gene expression signature of TREM2hi macrophages and γδ T cells predicts immunotherapy response. *Nat. Commun.* **11**, 5084 (2020).

96. Auslander, N. *et al.* Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat. Med.* **24**, 1545–1549 (2018).

97. Hugo, W. *et al.* Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016).

98. Bonavita, E. *et al.* Antagonistic inflammatory phenotypes dictate tumor fate and response to immune checkpoint blockade. *Immunity* **53**, 1215-1229.e8 (2020).

99. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**, 1550–1558 (2018).

100. Wang, L. *et al.* Enhancing KDM5A and TLR activity improves the response to immune checkpoint blockade. *Sci. Transl. Med.* **12**, eaax2282 (2020).

101. Davey, M. G., Hynes, S. O., Kerin, M. J., Miller, N. & Lowery, A. J. Ki-67 as a prognostic biomarker in invasive breast cancer. *Cancers* **13**, 4455 (2021).

102. Brandão, M., Pondé, N. & Piccart-Gebhart, M. Mammaprint™: a comprehensive review. *Future Oncol.* **15**, 207–224 (2019).

103. Slodkowska, E. A. & Ross, J. S. MammaPrint 70-gene signature: another milestone in personalized medical care for breast cancer patients. *Expert Rev. Mol. Diagn.* **9**, 417–422 (2009).

104. Syed, Y. Y. Oncotype DX breast recurrence score®: a review of its use in early-stage breast cancer. *Mol. Diagn. Ther.* **24**, 621–632 (2020).

105. Iwase, T. *et al.* A novel immunomodulatory 27-gene signature to predict response to neoadjuvant immunochemotherapy for primary triple-negative breast cancer. *Cancers* **13**, 4839 (2021).

106. Shen, J. *et al.* An 11-gene signature based on treatment responsiveness predicts radiation therapy survival benefit among breast cancer patients. *Front. Oncol.* **11**, 816053 (2021).

107. Mahoney, K. M., Freeman, G. J. & McDermott, D. F. The next immune-checkpoint inhibitors: PD-1/PD-L1 blockade in melanoma. *Clin. Ther.* **37**, 764–782 (2015).

108. Tsai, K. K., Zarzoso, I. & Daud, A. I. PD-1 and PD-L1 antibodies for melanoma. *Hum. Vaccines Immunother.* **10**, 3111–3116 (2014).

109. Wu, M. *et al.* Improvement of the anticancer efficacy of PD-1/PD-L1 blockade via combination therapy and PD-L1 regulation. *J. Hematol. Oncol.J Hematol Oncol* **15**, 24 (2022).

110. Lin, A. *et al.* Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Sci. Transl. Med.* **11**, eaaw8412 (2019).

111. Aguiar, P. N., De Mello, R. A., Hall, P., Tadokoro, H. & Lima Lopes, G. de. PD-L1 expression as a predictive biomarker in advanced non-small-cell lung cancer: updated survival data. *Immunotherapy* **9**, 499–506 (2017).

112. Clark, A. S. *et al.* Neoadjuvant T-DM1/pertuzumab and paclitaxel/trastuzumab/pertuzumab for HER2+ breast cancer in the adaptively randomized I-SPY2 trial. *Nat. Commun.* **12**, 6428 (2021).

113. Kim, Y., Nagy, M., Pollard, B. & Rajagopal, P. S. Optimizing transcriptome-based synthetic lethality predictions to improve precision oncology in early-stage breast cancer: BC-SELECT. Preprint at https://doi.org/10.1101/2024.08.15.608073 (2024).

114. Banna, G. L. *et al.* Are anti-PD1 and anti-PD-L1 alike? The non-small-cell lung cancer paradigm. *Oncol. Rev.* **14**, 490 (2020).

115. Sánchez-Magraner, L. *et al.* Functional engagement of the PD-1/PD-L1 complex but not PD-L1 expression is highly predictive of patient response to immunotherapy in non-small-cell lung cancer. *J. Clin. Oncol.* **41**, 2561–2570 (2023).

116. Zhao, Y., Liu, L. & Weng, L. Comparisons of underlying mechanisms, clinical efficacy and safety between anti-PD-1 and anti-PD-L1 immunotherapy: the state-of-the-art review and future perspectives. *Front. Pharmacol.* **12**, 714483 (2021).

117. Hoang, D.-T. *et al.* A deep-learning framework to predict cancer treatment response from histopathology images through imputed transcriptomics. *Nat. Cancer* **5**, 1305–1317 (2024).

118. Hoang, D.-T. *et al.* Path2Omics: Enhanced transcriptomic and methylation prediction accuracy from tumor histopathology. *BioRxiv Prepr. Serv. Biol.* 2025.02.26.640189 (2025) doi:10.1101/2025.02.26.640189.

119. Shulman, E. D. *et al.* AI-Driven Spatial Transcriptomics Unlocks Large-Scale Breast Cancer Biomarker Discovery from Histopathology. Preprint at https://doi.org/10.1101/2024.10.16.618609 (2024).

120. Stergiopoulos, S., Getz, K. A. & Blazynski, C. Evaluating the completeness of ClinicalTrials.gov. *Ther. Innov. Regul. Sci.* **53**, 307–317 (2019).

121. Clough, E. *et al*. NCBI GEO: archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res*. **52**, D138–D144 (2024).

122. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).

123. Langfelder, P. & Horvath, S. Fast R functions for robust correlations and hierarchical clustering. *J. Stat. Softw.* **46**, 1–17 (2012).

124. Tryka, K. A. *et al*. NCBI's database of genotypes and phenotypes: dbGaP. *Nucleic Acids Res*. **42**, D975–D979 (2014).

125. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

126. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2015).

127. Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).