Optimizing transcriptome-based synthetic lethality predictions to improve precision oncology in early-stage breast cancer: BC-SELECT

Yewon Kim[1],
Matthew Nagy[2],
Rebecca Pollard[1],
*Padma Sheila Rajagopal[1,3]

[1] Cancer and Data Science Laboratory, Center for Cancer Research, National Cancer Institute, Bethesda, MD
[2] Boston Children's Hospital, Boston, MD
[3] Women's Malignancies Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD

Abstract Word Count: 246 words
Word Count: 3,649 words
References: 75 references

Figure / Table Count:
Figures: 4
Tables: 2
Supplementary Figures: 2
Supplementary Tables: 3

Corresponding Author:
Padma Sheila Rajagopal, MD MPH MSc
Physician-Scientist Early Investigator

Address: Building 10, Room 4B54
Bethesda, MD 20892
Email: Sheila.rajagopal@nih.gov
Phone: 240-858-3169

RUNNING TITLE:
BC-SELECT: Treatment response prediction in early breast cancer

ABSTRACT

Patients with early-stage triple-negative (TNBC) or HER2-positive (HER2+) breast cancer do not have tools to predict response before treatment. Pathologic complete response (pCR) is a proxy that still requires neoadjuvant treatment. BC-SELECT is a clinically focused computational tool that leverages genetic interactions, such as synthetic lethality (SL) or rescue (SR), to predict treatment response from gene expression data. BC-SELECT builds on existing related tools (SELECT, ENLIGHT) to navigate competing treatment options.

BC-SELECT involves two main steps: (1) a “gene pair identification” step that identifies clinically relevant partner genes using *in vivo* and *in vitro* datasets for a drug target, and (2) “training/parameter tuning” using early-stage breast cancer clinical trials to develop predicted response scores to treatment. We evaluated BC-SELECT’s ability to predict responders to trastuzumab, poly (ADP-ribose) polymerase (PARP) inhibitors, and immunotherapy by training on three trials including 86 patients and validating on nine trials including 722 patients.

BC-SELECT significantly predicted treatment response in 6/9 early-stage trials, including in all PARP inhibitors (ROC-AUCs 0.6-0.8; Odds Ratio (OR): 1.9-3.5) and immunotherapy (ROC-AUCs 0.7-0.8; OR: 3.2-7.1). BC-SELECT distinguished between responders and non-responders in 8/9 trials (FDR-corrected Wilcoxon rank-sum test p-value <0.1). BC-SELECT outperformed SELECT and ENLIGHT in predicting

immunotherapy response. Specificity was comparable across subtypes, with HER2+ 0.64, TNBC 0.66, and hormone-receptor positive (HR+) 0.76.

There are no standard-of-care pre-treatment predictors of response in TNBC and HER2+ breast cancer. Our findings position BC-SELECT as an eventual decision-support tool to prioritize treatments approved for patients with early-stage breast cancer.

STATEMENT OF SIGNIFICANCE:

BC-SELECT is a clinically focused computational tool that leverages genetic interactions, such as synthetic lethality, to predict treatment response from gene expression data for PARP inhibitors and immunotherapy in early-stage breast cancer.

INTRODUCTION

Gene expression data (via RT-PCR-based panels) has transformed clinical management in early-stage breast cancer. However, this information is currently used overwhelmingly in patients with hormone receptor positive (HR+) disease, who benefit from the opportunity to undergo testing prior to receiving chemotherapy[1–9]. Patients with HER2-positive (HER2+) or triple-negative breast cancer (TNBC), unfortunately, currently undergo multiple neoadjuvant therapies and surgery to obtain pathologic complete response (pCR), a post-facto indicator of treatment response that then guides adjuvant treatment. There remains an ongoing need for clinically relevant predictive tools that can be applied before treatment to guide patients and oncology providers.

Many published computational approaches demonstrate the prognostic and potentially predictive potential of gene expression in early-stage TNBC or HER2+ breast cancer, but these approaches are not designed to navigate clinical challenges[10]. Recent studies support the use of “immune-related” gene signatures to predict immunotherapy response in breast cancer alongside clinical data[11–13]. HER2 gene expression is also evolving as a biomarker alongside immunohistochemistry and FISH testing[14,15]. Broader transcriptomic data can be routinely obtained on commercial testing and identifies not only fusion genes for FDA-approved therapeutic indications, but also gene expression levels, making this information ripe for integrating with existing clinical information[16,17].

Synthetic lethality (SL), a form of genetic interaction in which two genes must be concurrently inactivated to cause cell death, is a mechanism of action with known clinical relevance in breast cancer that can also be exploited through gene expression[18,19]. Poly (ADP-ribose) polymerase (PARP) inhibitors are approved to treat

early-stage and metastatic breast cancers with inherited *BRCA1/2* pathogenic mutations[20–22]. Other genetic interactions, such as synthetic dosage lethality (SDL), in which one gene must be over-expressed and another under-expressed[23], or synthetic rescue (SR), in which inactivation of one gene that would cause cell death is ameliorated by altered activation of another gene, are being actively studied for potential clinical applications[24].

Computational approaches to identify such genetic interactions from gene expression data have been recently developed. SELECT[25] uses large-scale cell-line and pan-cancer patient tumor transcriptomic data to identify clinically relevant SL/SR partners for a wide array of cancer therapies. A commercial adaptation, ENLIGHT[26], focuses on advanced metastatic cancers. Both algorithms predict treatment response in retrospective clinical trials[25,26]. However, despite the pan-cancer breadth of these tools, their application in early-stage breast cancer is limited. They were tested on breast cancer trials that included either endocrine therapy or trastuzumab, but neither on immunotherapy specifically in breast cancer nor on PARP inhibitors.

In this paper, we present BC-SELECT, an adaptation specifically designed to navigate a current area of need in clinical early-stage breast cancer by predicting response to multiple targeted therapies and immunotherapies. As proof of concept, we evaluated BC-SELECT's ability to predict responders to trastuzumab, PARP inhibitors, and immunotherapy by training on three trials with 86 patients and validating on nine additional trials with 722 patients. We demonstrated that BC-SELECT predicts treatment response in PARP inhibitors – building on prior work in SELECT and ENLIGHT – and immunotherapy across breast cancer subtypes, positioning it as a

potential tool for navigating the clinical challenge of competing approved therapeutic options.

RESULTS

Overview of the BC-SELECT framework and adaptations to predict pathologic complete response (pCR) to targeted therapy and immunotherapy in early-stage breast cancer

Figure 1 provides a visual overview of the BC-SELECT pipeline. BC-SELECT starts with two parts: “Gene pair identification” and “training / hyperparameter tuning.” To identify gene pairs, we used large-scale cell-line and breast cancer datasets to find clinically relevant, interacting gene pairs A and B (in which gene A is the target(s) for targeted therapy or immunotherapy, and multiple partner B genes are identified). We used three breast cancer clinical trials to train hyperparameters that distinguish responders from non-responders (Supplemental Table 1). Expression levels of partner B genes are used to generate a patient’s predicted response score to a treatment targeting gene A. We then validated BC-SELECT using nine separate, independent early-stage breast cancer clinical trials: Four using trastuzumab in HER2+ disease, three using PARP inhibitors (veliparib or olaparib) in HR+ or TNBC, and two using PD-1/PD-L1 immunotherapy (durvalumab and pembrolizumab) in HR+ or TNBC (Supplemental Table 1). All patients were analyzed only within their given clinical trial to limit batch effect issues; both microarray and RNA-seq data were used. Partner genes are listed in (Supplemental Table 2).

BC-SELECT significantly predicted responders and non-responders to PARP inhibitors and immunotherapy

Across all trials, responders (pCR achieved = 1) demonstrated consistently higher predicted scores from BC-SELECT compared to non-responders (pCR not achieved = 0).

We derived odds ratios between responders and non-responders based on the BC-SELECT final predicted response scores (Figure 2). Odds ratios were statistically significant in 6/9 trials using a threshold of 0.51, including all PARP inhibitor and all immunotherapy trials (Table 1). After correcting for multiple hypothesis testing, using the Wilcoxon rank-sum test with a corrected p-value <0.1, we also observed a statistically significant difference between final predicted response scores in responders and non-responders for both targeted therapy (6/7 trials) (Figure 3a) and immunotherapy (2/2 trials) (Figure 3b).

The final predicted response score used in BC-SELECT also outperformed the two components used to calculate the predicted response score: partner gene expression ratio and target gene expression (Table 2). Among targeted therapy options, BC-SELECT final predicted response scores were statistically significant in predicting pCR compared to chance (one-tailed empirical p-value < 0.025) in four trials (2/3 trials with a PARP inhibitor and 2/4 trastuzumab trials), with AUCs ranging from 0.63-0.8, while partner gene expression ratio and target gene expression were statistically significant each in two trials. For immunotherapy, BC-SELECT and partner gene expression ratio were both statistically significant in predicting response compared to chance in both trials, but ROC-AUCs were consistently higher with BC-SELECT (0.71 and 0.78 vs. 0.62 and 0.66). Target gene expression was significant in one of two trials.

Precision-recall curves for these trials, given the limitations of ROC-AUC alone, are shown in Supplementary Figure 1.

BC-SELECT performed better than SELECT and ENLIGHT in early-stage breast cancer immunotherapy trials

BC-SELECT achieved higher ROC-AUCs than SELECT and/or ENLIGHT in two immunotherapy cohorts (Figure 4). We were able to compare ENLIGHT's performance with durvalumab as this is in the pre-computed results generated by ENLIGHT, but pembrolizumab was not pre-computed. Neither algorithm previously published results with PARP inhibitors for the purposes of comparison.

BC-SELECT's specificity is comparable across subtypes

BC-SELECT was applied to predict responders and non-responders among treatments already approved for early-stage triple-negative breast cancer. As such, we were not seeking to determine which patients *should* or *should not* receive treatments – this is already clinically established – but rather, which patients may not experience as much benefit from these existing treatments (where they are less likely to respond). To this end, we assessed the specificity of the algorithm across breast cancer subtypes. Among 722 patients with known tissue-based subtypes (203 with HER2+, 172 with HR+, and 347 with TNBC), specificity was comparable, with HER2+ 0.64, HR+ 0.76, and TNBC 0.66.

DISCUSSION

We developed BC-SELECT, a computational algorithm that predicts responders to PARP inhibitors and immunotherapy in patients with early-stage breast cancer. This algorithm adds to predecessor tools focused on using synthetic lethality and gene expression to identify therapeutic options – SELECT and ENLIGHT – by successfully predicting PARP inhibitor therapy in breast cancer, deepening their prior coverage of breast cancer, and reinforcing the efficacy of the prior statistical design. BC-SELECT also builds on existing work in breast cancer evaluating gene expression for clinical prediction by evaluating multiple treatment options, providing a meaningful way to navigate the active clinical need of potential competing treatment recommendations.

BC-SELECT was developed with clinical utility in mind. We leveraged publicly available breast cancer-specific clinical trials and datasets using targeted therapy and immunotherapy. We added clinical covariates to the filtering of genetic interaction pairs and also filter gene partners through Nanostring gene lists. We use multiple measures of statistical significance to assess performance, including odds ratios, ROC-AUCs, precision-recall curves, and assessments of BC-SELECT and its score components in Monte-Carlo simulations to develop an estimate of performance vs. chance. Importantly, the algorithm is designed with future adaptation in mind, where additional datasets can be added to the training step, and additional clinical trials for other targeted therapies and immunotherapies can be added for tuning and validation.

BC-SELECT can be applied to multiple potential clinical contexts. The algorithm was developed to provide decision support in prioritizing among approved therapeutic options by prioritizing predicted responses using synthetic lethality. Patients with early-

stage TNBC due to pathogenic germline *BRCA1/2* variants, for example, may often be eligible for both PARP inhibitors and immunotherapy, with current guidelines providing limited evidence to support prioritization of one therapy versus another. Multiple clinical trials in metastatic disease (KEYLYNK-009, MEDIOLA, PEMBRACA) have failed to show an outcome benefit for combining these therapies[27–31], reflecting the need for specific clinical tools that can identify patients unlikely to respond to at least one of these agents. Another potential application would be to identify potential responders to therapy that are outside current clinical guidelines. For example, there has been an ongoing question of how to apply PARP inhibitors to patients without *BRCA1/2* mutations; BC-SELECT is well positioned to address this question in future by identifying other therapies may induce synthetic lethality with PARP inhibitors.

The overall performance of BC-SELECT, with its predictive AUC range of 0.6-0.8, is consistent with clinically used tests. Breast cancer polygenic risk scores achieve AUCs in the 0.6 range, and Oncotype / other breast cancer gene expression assays AUCs near 0.75[32,33]. As such, BC-SELECT merits ongoing work for translation.

The partner genes selected by BC-SELECT reflect known biology. Five of the 10 partner genes prioritized for immunotherapy demonstrate potential efficacy in studies involving synthetic lethality and the immune system. *CEP55, TTK,* and *CCND3* depletion/inhibition with other compounds can lead to breast cancer cell death *in vitro*[34–37]. Combination immunotherapies with *LAG3* inhibitors performed better than *LAG3* inhibitors alone[38,39] in early-phase clinical trials. *TMEM173* (STING) also influences synthetic lethality via an immune response[40]. Among the partner genes prioritized for PARP inhibitors, the gene family most represented were the solute carriers (*SLC16A6,*

SLC27A4, SLC4A2, SLC7A14, and *SLC7A8*). *SLC7A11* has been implicated in ferroptosis and joint mechanism of response with PARP inhibitors where *BRCA1/2* remain proficient[41]. Most of the other prioritized genes are cell cycle pathway members, consistent with the known mechanism of action for PARP inhibitor synthetic lethality[42].

Other published predictive gene expression methods are either not trained as robustly in breast cancer (SELECT[25], ENLIGHT[26]) or do not incorporate genetic interactions or multiple treatment options (immune gene signatures for immunotherapy). It is possible that BC-SELECT's improved performance relative to SELECT and ENLIGHT is largely due to training on tumor-specific data. TNBC subtyping is a separate translational approach that uses gene expression data to classify breast cancers within TNBCs and identify potential treatment targets[43]. This approach is not competitive to ours, but rather, complementary, given that our approach adds responses based on predicted genetic interactions.

By incorporating target gene expression, BC-SELECT still reinforces prior research on the predictive role of gene expression. Target gene expression, specifically the product of *PDCD1* and *CD274*, performed well in predicting response to immunotherapy. *PD-L1* is an existing therapeutic biomarker via protein expression, with an ongoing challenge being the best way to optimize its clinical deployment[44]. RNA-based expression of *PD-L1* and *CD274* are known to correlate with immunotherapy responders[45–47], and incorporated into predictive immune signatures that, though not used by themselves for treatment response prediction[12]. Our results highlight the benefit of including the information in the final predicted response score generated by BC-SELECT over target gene expression alone.

BC-SELECT has limitations consistent with SELECT and ENLIGHT. It requires specific drug target(s) and can only be applied to targeted therapy and immunotherapy. Many chemotherapies, unfortunately, have numerous target effects by mechanism, limiting the ability to predict chemotherapy response using this tool. BC-SELECT also does not demonstrate as effective prediction if the mechanism of action is not driven by synthetic lethality, as we observe with trastuzumab in this study. We cannot confirm results in treatments for which we did not train. We describe our testing of only three therapies, but as more clinical trials in breast cancer offer publicly available data, BC-SELECT's design allows this data to be incorporated in future into the training and validation.

There are restrictions to the trials we could include for this study. We observed decreased prediction accuracy with clinical trials featuring fewer than 25 patients. We were not able to include trials that only generated gene expression for a small number of genes, making it difficult to incorporate older consortia trials. The small number of trials with RNA-seq (n=2) as opposed to microarray data (n=7) also prevents us from drawing conclusions in the study about the impact if any on data type.

In summary, we present BC-SELECT, a computational algorithm to predict responders to PARP inhibitors and immunotherapy in patients with early-stage breast cancer. Given the increased success rates of oncology trials studying synthetic lethality outside of PARP inhibitors[48], it is worth exploring genetic interactions such as synthetic lethality and applications that can use these interactions to select treatments. BC-SELECT demonstrates predictive performance in the context of competing approved

therapeutic options, enabling the differentiation of responders and non-responders for future translational purposes and eventual clinical decision support.

METHODS

Gene pair identification

Figure 1 shows the workflow of BC-SELECT. In BC-SELECT, we first identify pairs of genes that interact via synthetic lethality (SL), synthetic dosage lethality (SDL), or synthetic rescue (SR), where one of the genes is a drug target. For targeted therapies, we test for SL/SDL gene pairs. DrugBank is the reference for drug target genes[49]. For immunotherapy, we test for SR pairs. *PDCD1* and *CD274* were our drug targets. We focus on targeting PD1/PDL1, as other classes of immunotherapy are not yet approved in breast cancer and did not have corresponding clinical trial data available.

Gene pairs that pass the following three filtering steps are used by BC-SELECT to predict treatment response:

(1) A **large-scale gene-pair screen using *in vitro* tumor cell-line data or *in vivo* tumor data.** For *targeted therapy*, we perform a Wilcoxon rank-sum test to identify pairs of genes from published studies on genetic dependencies / synthetic lethality screening, from databases such as DepMap and the Cancer Cell Line Encyclopedia[50–55], and published studies on genetic relationships to pharmacology [56–58]. For *immunotherapy*, we perform a hypergeometric test[59] to identify pairs of genes observed as under-expressed more frequently than expected by chance in three publicly available datasets of gene expression data (either microarray or normalized RNA-seq) from patients with breast cancer: TCGA (n=1075)[60], METABRIC (n=1989)[61], and SCAN-B (n=2969)[62]. Each dataset is separately assessed to avoid batch effect. Significant gene pairs are

advanced to step 2, where statistical significance is determined using a pre-specified FDR-corrected threshold of p=0.2 to avoid overly conservative exclusion of potentially relevant genetic interactions.

(2) **Cox proportional hazards models** that include the genetic interaction pairs identified in step 1. Clinical variables are included as covariates in addition to gene expression values to determine if gene expression remains significantly associated with patient survival in a clinical context. To incorporate gene expression into a Cox proportional hazards model, we split expression for a gene into tertiles (upregulated in the top tertile, “no change” in the middle tertile, and downregulated in the bottom tertile). We incorporate cancer-specific covariates, including age, stage, sex, and race as these are available in each dataset.

For targeted therapy: We perform six separate Cox proportional hazard analyses: One for SL gene pairs and one for SDL gene pairs, each performed separately in the TCGA[60], METABRIC[61], and SCAN-B[62] datasets to avoid batch effect confounding.

For immunotherapy: We perform three separate Cox proportional hazard analyses using the TCGA[60], METABRIC[61], and SCAN-B[62] datasets and SR gene pairs.

Significance of a given candidate gene pair within at least one Cox proportional hazard model is determined by using the multiplicity-corrected marginal p-value of covariates for gene expression and the p-value from the associated log-likelihood ratio test. If a gene pair has multiple significant p-values, the minimum p-value across these analyses is taken as the representative value for that gene. Again, significance is determined by a pre-specified FDR-corrected threshold of p=0.2 to avoid overly conservative gene pair exclusion. Gene pairs significantly associated with survival through at least one Cox proportional hazards analysis are advanced to step 3.

(3) **Evolutionary analysis of preserved gene expression relationships** using the Euclidean distance of phylogenetic similarity scores across multiple species [63] and feature weight [64].

For *immunotherapy:* We compute the phylogenetic similarity score between both *PDCD1* and *CD274* and an associated candidate partner gene. The final phylogenetic score for the partner gene is determined as the smaller value between the two similarity scores generated for *PDCD1* and *CD274*.

We identify phylogenetically linked candidate partner genes using a threshold (acceptable maximum Euclidean distance between two linked genes) of 10.5 for targeted therapy[59] and 23.15 for immunotherapy[25]. Significant gene pairs from this step are then used to predict treatment response.

Training / Parameter tuning

We trained the hyperparameters of treatment response prediction using three breast cancer clinical trials: GSE66399[65], GSE160568[66], and NCT02489448[67] (Supplementary Table 1). This approach allows BC-SELECT to avoid overfitting issues while achieving flexibility across trials. We performed a grid search to optimize the hyperparameters of the number of significant gene pairs required for prediction of treatment response, and the odds ratio threshold to define responders/non-responders.

For *targeted therapy*, we select the top 55 SL/SDL gene pairs, with prioritization based on adjusted significance in Cox proportional hazards testing. We used only either SL or SDL depending on predictive performance. For *immunotherapy*, we selected the top 10 SR gene pairs, with prioritization based on phylogenetic similarity scores. The odds ratio threshold of greater than 0.51 was used to differentiate between responders and non-responders.

BC-SELECT predicted response score

The predicted response score in BC-SELECT is the product of the **partner gene expression ratio (a/b/c/d)** and the **drug target gene expression level**.

Partner gene expression ratio:

The numerator is a ratio of:

- SL: Down-regulated partner genes OR

  SDL: Up-regulated partner genes over all 55 SDL gene pairs identified for a *targeted therapy*
- SR: 1 minus the ratio of down-regulated partner genes over all 10 SR gene pairs identified for *immunotherapy* (as we are measuring the likelihood of drug response)

The denominator is a ratio of:

- SL: Down-regulated genes OR

  SDL: Up-regulated genes among 776 breast cancer-specific genes for *targeted therapy*
- SR: 1 minus the ratio of down-regulated genes among 1154 immune-specific genes for *immunotherapy*

The breast cancer-specific and immune-specific gene lists were selected through commercial Nanostring panels.[68,69] The two Nanostring gene lists are provided in Supplementary Table 3.

Drug target gene expression level

This is the gene expression level of the drug target in a given gene pair, as measured via either microarray or normalized RNA-seq. Trials or datasets were not combined to avoid batch effect correction issues. For drugs with multiple targets, we use the mean expression of all target genes in targeted therapies. For immunotherapy, we calculate the simple product of *PDCD1* and *CD274* expression.

Validation trials

We used publicly available clinical trials with at least 25 patients with early-stage breast cancer, individual level gene expression data, and reported treatment response. We selected pCR to measure response, as this the most commonly studied trial treatment response variable. This was reported across studies as a binary measure with either 0 (residual disease) or 1 (complete response). We included a total of 9 clinical trials, with 7 across 3 targeted therapies (trastuzumab, olaparib, veliparib)[13,70–75] and 2 clinical trials for immunotherapy[13,75]. Trials used in this paper are described in Supplementary Table 1.

Validation of treatment response prediction

We compared the predicted response scores of BC-SELECT for a given therapy in a given patient to the actual responses reported in each clinical trial. We calculated odds ratios with corresponding 95% confidence intervals and p-values using Fisher's exact test, as well as ROC-AUC and precision-recall curves to assess prediction performance.

Changes relative to SELECT

We adapted the original code from SELECT[25] as follows:

- We remove copy number data to facilitate future addition of publicly available breast cancer datasets.
- We also add additional clinical covariates to the survival analyses in the Cox proportional hazards filtering step during the gene pair identification step to improve clinical utility.

- For targeted therapies, we prioritize candidate partner genes based on their significance in the survival analysis filtering step.
- We introduce SDL as a separate concept and score, wherein a higher SDL score signifies a higher likelihood of cancer cell death and a stronger drug response.
- We have revised the calculation of the final predicted response score, including the incorporation of drug target gene expression and the partner gene expression ratio.
- We perform Monte Carlo simulations to generate empirical p-values to assess the performance of BC-SELECT relative to chance.

Statistical analysis

Statistical tests used to prioritize partner genes were described in "**Partner gene identification".**

We estimated the statistical significance of BC-SELECT's predicted response score and components compared to chance. We used Monte-Carlo simulations to develop a "null distribution of AUCs" for each trial. Using 55 randomly selected gene pairs for targeted therapy and 10 randomly selected gene pairs for immunotherapy, we calculated ROC-AUC values for these randomly selected genes and iterated this process 5000 times to yield the "null distribution". We then compared the BC-SELECT's predictive ROC-AUC to those achieved by chance. This process produced empirical p-values, in which the top 2.5% (statistical significance threshold p=0.025), corresponds to the single tail with

predicted response better than two standard deviations from the mean in the null distribution (Supplementary Figure 2).

All analyses were conducted using R v.4.2. To control the false discovery rate, we apply the Benjamini–Hochberg procedure (BH step-up procedure) as a multiplicity correction.

References

1. Prat A, Ellis MJ, Perou CM. Practical implications of gene-expression-based assays for breast oncologists. *Nat Rev Clin Oncol*. 2011;9(1):48-57. doi:10.1038/nrclinonc.2011.178

2. Kalinsky K, Barlow WE, Gralow JR, et al. 21-Gene Assay to Inform Chemotherapy Benefit in Node-Positive Breast Cancer. *N Engl J Med*. 2021;385(25):2336-2347. doi:10.1056/NEJMoa2108873

3. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med*. 2016;375(8):717-729. doi:10.1056/NEJMoa1602253

4. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med*. 2018;379(2):111-121. doi:10.1056/NEJMoa1804710

5. Piccart M, van 't Veer LJ, Poncet C, et al. 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. *Lancet Oncol*. 2021;22(4):476-488. doi:10.1016/S1470-2045(21)00007-3

6. Nielsen T, Wallden B, Schaper C, et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer*. 2014;14:177. doi:10.1186/1471-2407-14-177

7. Dubsky P, Brase JC, Jakesz R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. *Br J Cancer*. 2013;109(12):2959-2964. doi:10.1038/bjc.2013.671

8. O'Regan RM, Zhang Y, Fleming GF, et al. Breast Cancer Index in Premenopausal Women With Early-Stage Hormone Receptor-Positive Breast Cancer. *JAMA Oncol*. Published online August 15, 2024:e243044. doi:10.1001/jamaoncol.2024.3044

9. Bottosso M, Miglietta F, Vernaci GM, et al. Gene Expression Assays to Tailor Adjuvant Endocrine Therapy for HR+/HER2- Breast Cancer. *Clin Cancer Res*. 2024;30(14):2884-2894. doi:10.1158/1078-0432.CCR-23-4020

10. Lehmann BD, Colaprico A, Silva TC, et al. Multi-omics analysis identifies therapeutic vulnerabilities in triple-negative breast cancer subtypes. *Nat Commun*. 2021;12(1):6276. doi:10.1038/s41467-021-26502-6

11. Safonov A, Jiang T, Bianchini G, et al. Immune Gene Expression Is Associated with Genomic Aberrations in Breast Cancer. *Cancer Res*. 2017;77(12):3317-3324. doi:10.1158/0008-5472.CAN-16-3478

12. Fernandez-Martinez A, Pascual T, Singh B, et al. Prognostic and Predictive Value of Immune-Related Gene Expression Signatures vs Tumor-Infiltrating Lymphocytes in Early-Stage ERBB2/HER2-Positive Breast Cancer: A Correlative Analysis of the CALGB 40601 and PAMELA Trials. *JAMA Oncol*. 2023;9(4):490-499. doi:10.1001/jamaoncol.2022.6288

13. Wolf DM, Yau C, Wulfkuhle J, et al. Redefining breast cancer subtypes to guide treatment prioritization and maximize response: Predictive biomarkers across 10 cancer therapies. *Cancer Cell*. 2022;40(6):609-623.e6. doi:10.1016/j.ccell.2022.05.005

14. Verschoor N, Smid M, Jager A, Sleijfer S, Wilting SM, Martens JWM. Integrative whole-genome and transcriptome analysis of HER2-amplified metastatic breast cancer. *Breast Cancer Res*. 2023;25(1):145. doi:10.1186/s13058-023-01743-z

15. Stewart Mark. *Digital and Computational Pathology Tool Harmonization PATH Project Session 1 Presentation*. Digital PATH Project Working Group, Friends of Cancer Research https://friendsofcancerresearch.org/wp-content/uploads/Digital-and-Computational-Pathology-Tool-Harmonization-PATH-Project-Session-1-Presentation.pdf

16. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®) Breast Cancer v.4.2024. https://www.nccn.org/professionals/physician_gls/pdf/breast.pdf

17. Priedigkeit N, Lebrón-Torres A, Liao J, et al. Abstract GS03-09: Characterization and proposed therapeutic exploitation of fusion RNAs in metastatic breast cancers. *Cancer Research*. 2024;84(9_Supplement):GS03-09. doi:10.1158/1538-7445.SABCS23-GS03-09

18. Ashworth A, Lord CJ. Synthetic lethal therapies for cancer: what's next after PARP inhibitors? *Nat Rev Clin Oncol*. 2018;15(9):564-576. doi:10.1038/s41571-018-0055-6

19. O'Neil NJ, Bailey ML, Hieter P. Synthetic lethality and cancer. *Nat Rev Genet*. 2017;18(10):613-623. doi:10.1038/nrg.2017.47

20. Litton JK, Rugo HS, Ettl J, et al. Talazoparib in Patients with Advanced Breast Cancer and a Germline BRCA Mutation. *N Engl J Med*. 2018;379(8):753-763. doi:10.1056/NEJMoa1802905

21. Robson M, Im SA, Senkus E, et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med*. 2017;377(6):523-533. doi:10.1056/NEJMoa1706450

22. Tutt ANJ, Garber JE, Kaufman B, et al. Adjuvant Olaparib for Patients with BRCA1- or BRCA2-Mutated Breast Cancer. *N Engl J Med*. 2021;384(25):2394-2405. doi:10.1056/NEJMoa2105215

23. Megchelenbrink W, Katzir R, Lu X, Ruppin E, Notebaart RA. Synthetic dosage lethality in the human metabolic network is highly predictive of tumor growth and cancer patient survival. *Proc Natl Acad Sci U S A*. 2015;112(39):12217-12222. doi:10.1073/pnas.1508573112

24. Sahu AD, S Lee J, Wang Z, et al. Genome-wide prediction of synthetic rescue mediators of resistance to targeted and immunotherapy. *Mol Syst Biol*. 2019;15(3):e8323. doi:10.15252/msb.20188323

25. Lee JS, Nair NU, Dinstag G, et al. Synthetic lethality-mediated precision oncology via the tumor transcriptome. *Cell*. 2021;184(9):2487-2502.e13. doi:10.1016/j.cell.2021.03.030

26. Dinstag G, Shulman ED, Elis E, et al. Clinically oriented prediction of patient response to targeted and immunotherapies from the tumor transcriptome. *Med*. 2023;4(1):15-30.e8. doi:10.1016/j.medj.2022.11.001

27. Cussac AL, Rugo HS, Robson ME, et al. 198P Pembrolizumab plus olaparib vs pembrolizumab plus chemotherapy after induction with pembrolizumab plus chemotherapy for locally recurrent inoperable or metastatic TNBC: Patient-reported outcomes from KEYLYNK-009. *ESMO Open*. 2024;9:103220. doi:10.1016/j.esmoop.2024.103220

28. Thawani R, Kummar S. Combining PARP Inhibitor With Immunotherapy-Does the Promise of Preclinical Data Translate to Clinic? *JAMA Oncol*. 2023;9(1):25-27. doi:10.1001/jamaoncol.2022.4591

29. Domchek SM, Postel-Vinay S, Im SA, et al. Olaparib and durvalumab in patients with germline BRCA-mutated metastatic breast cancer (MEDIOLA): an open-label, multicentre, phase 1/2, basket study. *Lancet Oncol*. 2020;21(9):1155-1164. doi:10.1016/S1470-2045(20)30324-7

30. Cortesi L, Venturelli M, Cortesi G, et al. A phase II study of pembrolizumab plus carboplatin in BRCA-related metastatic breast cancer (PEMBRACA). *ESMO Open*. 2023;8(2):101207. doi:10.1016/j.esmoop.2023.101207

31. Maintenance Pembrolizumab/Olaparib Misses Survival End Points in Locally Advanced or Metastatic TNBC. OncLive. December 6, 2023. Accessed September 1, 2024. https://www.onclive.com/view/maintenance-pembrolizumab-olaparib-misses-survival-end-points-in-locally-advanced-or-metastatic-tnbc

32. Zirpoli GR, Pfeiffer RM, Bertrand KA, Huo D, Lunetta KL, Palmer JR. Addition of polygenic risk score to a risk calculator for prediction of breast cancer in US Black women. *Breast Cancer Res*. 2024;26(1):2. doi:10.1186/s13058-023-01748-8

33. de Hond AAH, Steyerberg EW, van Calster B. Interpreting area under the receiver operating characteristic curve. *Lancet Digit Health*. 2022;4(12):e853-e855. doi:10.1016/S2589-7500(22)00188-1

34. Wang G, Chen B, Su Y, Qu N, Zhou D, Zhou W. CEP55 as a Promising Immune Intervention Marker to Regulate Tumor Progression: A Pan-Cancer Analysis with Experimental Verification. *Cells*. 2023;12(20):2457. doi:10.3390/cells12202457

35. Zhu D, Xu S, Deyanat-Yazdi G, et al. Synthetic Lethal Strategy Identifies a Potent and Selective TTK and CLK1/2 Inhibitor for Treatment of Triple-Negative Breast Cancer with a Compromised G1-S Checkpoint. *Mol Cancer Ther*. 2018;17(8):1727-1738. doi:10.1158/1535-7163.MCT-17-1084

36. Martins MM, Zhou AY, Corella A, et al. Linking tumor mutations to drug responses via a quantitative chemical-genetic interaction map. *Cancer Discov*. 2015;5(2):154-167. doi:10.1158/2159-8290.CD-14-0552

37. Sinha D, Kalimutho M, Lopez AJ, Khanna KK. Synergistic inhibition of CEP55 induces mitotic catastrophe and specifically targets aggressive breast cancer. *Annals of Oncology*. 2017;28:v583. doi:10.1093/annonc/mdx390.033

38. LAG-3 transcriptomic expression patterns across malignancies: Implications for precision immunotherapeutics - Adashek - 2023 - Cancer Medicine - Wiley Online Library. Accessed February 19, 2025. https://onlinelibrary.wiley.com/doi/full/10.1002/cam4.6000

39. Clinical blockade of PD1 and LAG3 — potential mechanisms of action | Nature Reviews Immunology. Accessed February 19, 2025. https://www.nature.com/articles/nri3790

40. STING agonism enhances anti-tumor immune responses and therapeutic efficacy of PARP inhibition in BRCA-associated breast cancer | npj Breast Cancer. Accessed February 19, 2025. https://www.nature.com/articles/s41523-022-00471-5

41. Hong T, Lei G, Chen X, et al. PARP inhibition promotes ferroptosis via repressing SLC7A11 and synergizes with ferroptosis inducers in BRCA-proficient ovarian cancer. *Redox Biology*. 2021;42:101928. doi:10.1016/j.redox.2021.101928

42. Synthetic lethality prediction in DNA damage repair, chromatin remodeling and the cell cycle using multi-omics data from cell lines and patients. | Scientific Reports. Accessed February 19, 2025. https://www.nature.com/articles/s41598-023-34161-4

43. Lehmann BD, Pietenpol JA, Tan AR. Triple-negative breast cancer: molecular subtypes and new targets for therapy. *Am Soc Clin Oncol Educ Book*. Published online 2015:e31-39. doi:10.14694/EdBook_AM.2015.35.e31

44. Grossman JE, Vasudevan D, Joyce CE, Hildago M. Is PD-L1 a consistent biomarker for anti-PD-1 therapy? The model of balstilimab in a virally-driven tumor. *Oncogene*. 2021;40(8):1393-1395. doi:10.1038/s41388-020-01611-6

45. Conroy JM, Pabla S, Nesline MK, et al. Next generation sequencing of PD-L1 for predicting response to immune checkpoint inhibitors. *J Immunother Cancer*. 2019;7(1):18. doi:10.1186/s40425-018-0489-5

46. Kang SY, Heo YJ, Kwon GY, Kim KM. Expression of CD274 mRNA Measured by qRT-PCR Correlates With PD-L1 Immunohistochemistry in Gastric and Urothelial Carcinoma. *Front Oncol*. 2022;12:856444. doi:10.3389/fonc.2022.856444

47. Paré L, Pascual T, Seguí E, et al. Association between PD1 mRNA and response to anti-PD1 monotherapy across multiple cancer types. *Ann Oncol*. 2018;29(10):2121-2128. doi:10.1093/annonc/mdy335

48. Schäffer AA, Chung Y, Kammula AV, Ruppin E, Lee JS. A systematic analysis of the landscape of synthetic lethality-driven precision oncology. *Med*. 2024;5(1):73-89.e9. doi:10.1016/j.medj.2023.12.009

49. Knox C, Wilson M, Klinger CM, et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. *Nucleic Acids Res*. 2024;52(D1):D1265-D1275. doi:10.1093/nar/gkad976

50. Aguirre AJ, Meyers RM, Weir BA, et al. Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting. *Cancer Discov*. 2016;6(8):914-929. doi:10.1158/2159-8290.CD-16-0154

51. Cheung HW, Cowley GS, Weir BA, et al. Systematic investigation of genetic vulnerabilities across cancer cell lines reveals lineage-specific dependencies in ovarian cancer. *Proc Natl Acad Sci U S A*. 2011;108(30):12372-12377. doi:10.1073/pnas.1109363108

52. Cowley GS, Weir BA, Vazquez F, et al. Parallel genome-scale loss of function screens in 216 cancer cell lines for the identification of context-specific genetic dependencies. *Sci Data*. 2014;1:140035. doi:10.1038/sdata.2014.35

53. Marcotte R, Brown KR, Suarez F, et al. Essential gene profiles in breast, pancreatic, and ovarian cancer cells. *Cancer Discov*. 2012;2(2):172-189. doi:10.1158/2159-8290.CD-11-0224

54. Marcotte R, Sayad A, Brown KR, et al. Functional Genomic Landscape of Human Breast Cancer Drivers, Vulnerabilities, and Resistance. *Cell*. 2016;164(1-2):293-309. doi:10.1016/j.cell.2015.11.062

55. Tsherniak A, Vazquez F, Montgomery PG, et al. Defining a Cancer Dependency Map. *Cell*. 2017;170(3):564-576.e16. doi:10.1016/j.cell.2017.06.010

56. Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483(7391):603-607. doi:10.1038/nature11003

57. Basu A, Bodycombe NE, Cheah JH, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell*. 2013;154(5):1151-1161. doi:10.1016/j.cell.2013.08.003

58. Iorio F, Knijnenburg TA, Vis DJ, et al. A Landscape of Pharmacogenomic Interactions in Cancer. *Cell*. 2016;166(3):740-754. doi:10.1016/j.cell.2016.06.017

59. Lee JS, Das A, Jerby-Arnon L, et al. Harnessing synthetic lethality to predict the response to cancer treatment. *Nat Commun*. 2018;9(1):2546. doi:10.1038/s41467-018-04647-1

60. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013;45(10):1113-1120. doi:10.1038/ng.2764

61. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403):346-352. doi:10.1038/nature10983

62. Brueffer C, Vallon-Christersson J, Grabau D, et al. Clinical Value of RNA Sequencing-Based Classifiers for Prediction of the Five Conventional Breast Cancer Biomarkers: A Report From the Population-Based Multicenter Sweden Cancerome Analysis Network-Breast Initiative. *JCO Precis Oncol*. 2018;2:PO.17.00135. doi:10.1200/PO.17.00135

63. Tabach Y, Golan T, Hernández-Hernández A, et al. Human disease locus discovery and mapping to molecular pathways through phylogenetic profiling. *Mol Syst Biol*. 2013;9:692. doi:10.1038/msb.2013.50

64. Martin FJ, Amode MR, Aneja A, et al. Ensembl 2023. *Nucleic Acids Res*. 2023;51(D1):D933-D941. doi:10.1093/nar/gkac958

65. Guarneri V, Dieci MV, Frassoldati A, et al. Prospective Biomarker Analysis of the Randomized CHER-LOB Study Evaluating the Dual Anti-HER2 Treatment With Trastuzumab and Lapatinib Plus Chemotherapy as Neoadjuvant Therapy for HER2-Positive Breast Cancer. *Oncologist*. 2015;20(9):1001-1010. doi:10.1634/theoncologist.2015-0138

66. Liu X, Ge Z, Yang F, et al. Identification of biomarkers of response to preoperative talazoparib monotherapy in treatment naïve gBRCA+ breast cancers. *NPJ Breast Cancer*. 2022;8(1):64. doi:10.1038/s41523-022-00427-9

67. Foldi J, Silber A, Reisenbichler E, et al. Neoadjuvant durvalumab plus weekly nab-paclitaxel and dose-dense doxorubicin/cyclophosphamide in triple-negative breast cancer. *NPJ Breast Cancer*. 2021;7(1):9. doi:10.1038/s41523-021-00219-7

68. nCounter Breast Cancer 360. NanoString. 2024. Accessed July 8, 2024. https://nanostring.com/products/ncounter-assays-panels/oncology/breast-cancer-360/

69. PanCancer Immune Profiling. NanoString. 2024. Accessed July 8, 2024. https://nanostring.com/products/ncounter-assays-panels/oncology/pancancer-immune-profiling/

70. Prat A, Bianchini G, Thomas M, et al. Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. *Clin Cancer Res*. 2014;20(2):511-521. doi:10.1158/1078-0432.CCR-13-0239

71. Liu JC, Voisin V, Bader GD, et al. Seventeen-gene signature from enriched Her2/Neu mammary tumor-initiating cells predicts clinical outcome for human HER2+:ERα- breast cancer. *Proc Natl Acad Sci U S A*. 2012;109(15):5832-5837. doi:10.1073/pnas.1201105109

72. Shen K, Qi Y, Song N, et al. Cell line derived multi-gene predictor of pathologic response to neoadjuvant chemotherapy in breast cancer: a validation study on US Oncology 02-103 clinical trial. *BMC Med Genomics*. 2012;5:51. doi:10.1186/1755-8794-5-51

73. Sammut SJ, Crispin-Ortuzar M, Chin SF, et al. Multi-omic machine learning predictor of breast cancer therapy response. *Nature*. 2022;601(7894):623-629. doi:10.1038/s41586-021-04278-5

74. Loibl S, O'Shaughnessy J, Untch M, et al. Addition of the PARP inhibitor veliparib plus carboplatin or carboplatin alone to standard neoadjuvant chemotherapy in triple-negative breast cancer (BrighTNess): a randomised, phase 3 trial. *Lancet Oncol*. 2018;19(4):497-509. doi:10.1016/S1470-2045(18)30111-6

75. Pusztai L, Yau C, Wolf DM, et al. Durvalumab with olaparib and paclitaxel for high-risk HER2-negative stage II/III breast cancer: Results from the adaptively randomized I-SPY2 trial. *Cancer Cell*. 2021;39(7):989-998.e5. doi:10.1016/j.ccell.2021.05.009

Conflicts of Interest:

Yewon Kim: None
Matthew Nagy: None
Becca Pollard: None
Padma Sheila Rajagopal: None

Acknowledgments:

We wish to thank Dr. Eytan Ruppin, Dr. Lipika Ray, and Dr. Nishanth Nair for their support and counsel throughout the drafting of this manuscript. We also wish to thank Dr. Kim Blenman, Dr. Lajos Pusztai and Dr. Daniel Stover for their generosity in providing individual-level training data or data-use recommendations in support of this project. This work utilized the computational resources of the NIH HPC Biowulf cluster. (http://hpc.nih.gov)

Data availability:

The data analyzed in this study is from publicly available repositories as described in the Supplementary Files. Source code for BC-SELECT is available on GitHub: https://github.com/yewon-kim76/BC-SELECT, and the list of references can be found in Zotero: https://www.zotero.org/yewon76/collections/G396YYNB.

Author contributions:

Yewon Kim: Methodology, Software, Validation, Analysis, Writing - Original And Revision. **Becca Pollard**: Data Curation. **Matthew Nagy**: Software, Validation, Writing – Revision. **Padma Sheila Rajagopal**: Conceptualization, Methodology, Validation, Writing - Revision, Resources, Data Curation, Supervision, Administration, Funding.

Tables and Figures

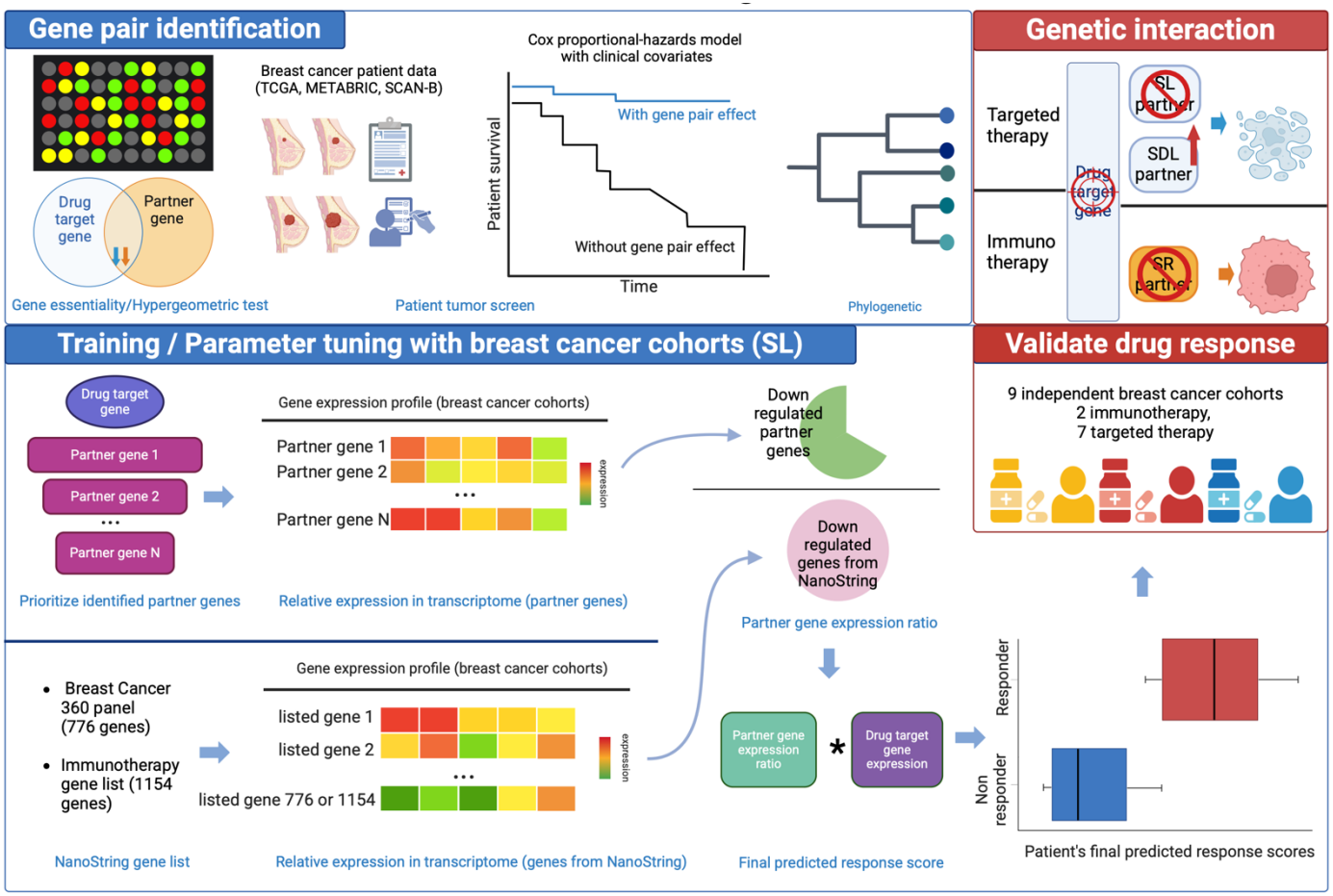


Figure 1: A graphical overview illustrating the adaptation of BC-SELECT pipeline. BC-SELECT involves two steps (in blue): gene pair identification and hyperparameter tuning. The first step focuses on identifying gene pairs involved in genetic interactions. We then adjust hyperparameters to optimize drug response prediction performance during the second step. We are then able to apply this to unseen clinical trials and did so in nine validation breast cancer clinical trials.

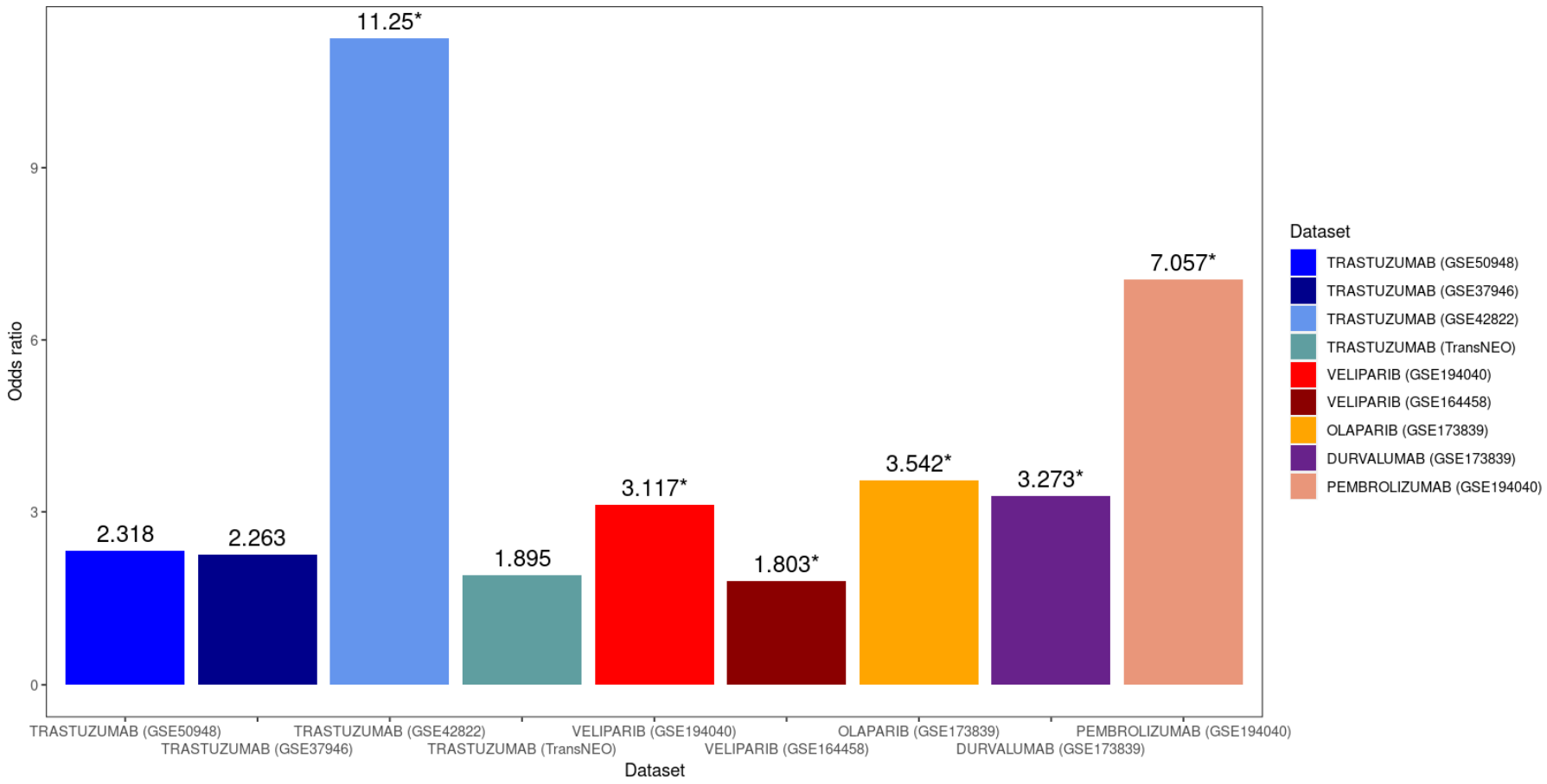
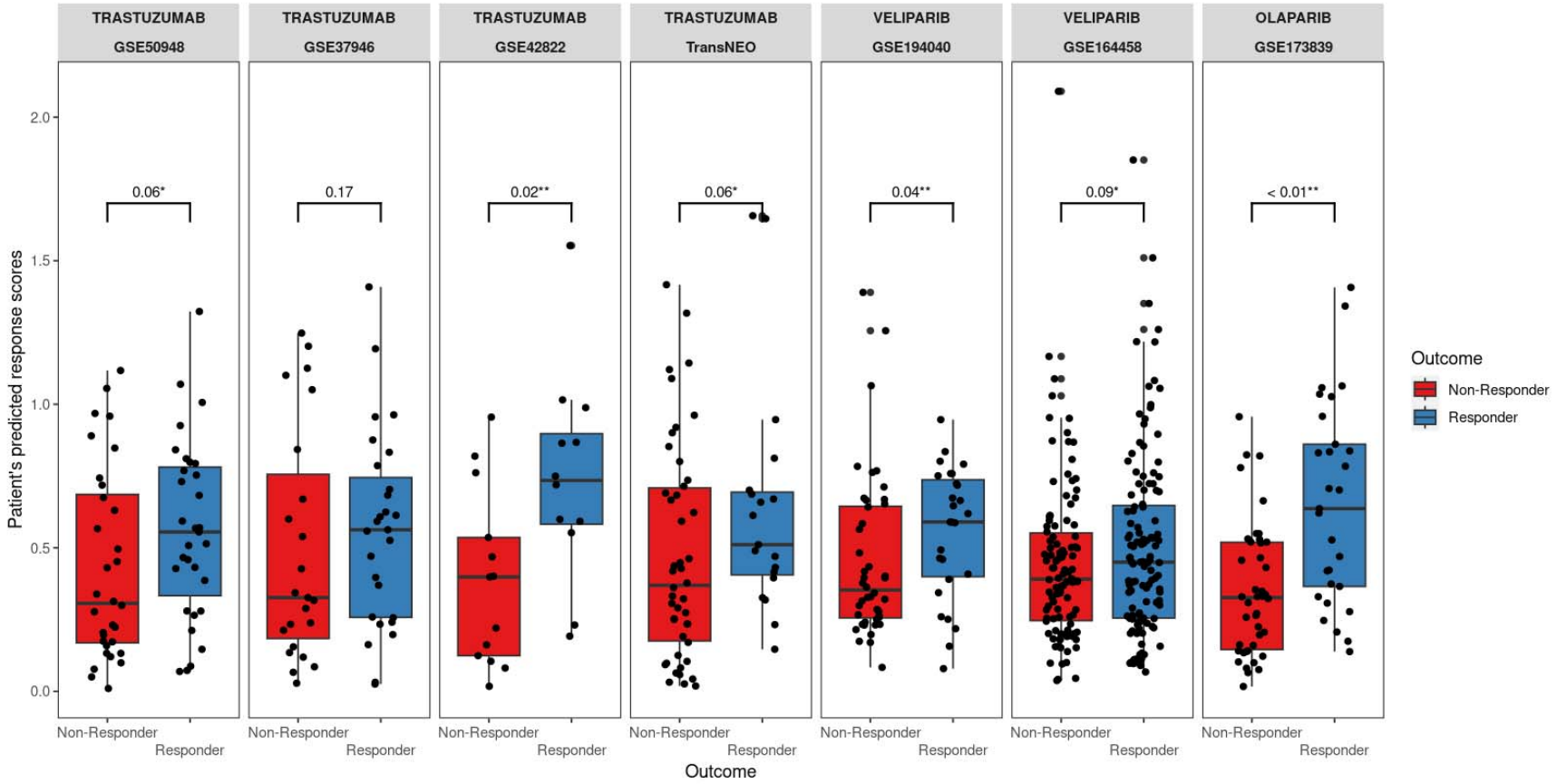


Figure 2: Odds ratios between responders and non-responders using final predicted response scores from BC-SELECT. X-axis shows the specific clinical trial, and Y-axis shows the odds ratio. Using a threshold of 0.51, we calculated the odds ratios between responders and non-responders across 9 breast cancer clinical trials. Trastuzumab is shown in blue/green, PARP inhibitors in red/orange, and immunotherapy in purple/pink. All odds ratios were greater than 1, and prediction in 6 trials showed statistical significance (marked with *), with the 95% confidence interval not containing 1. Detailed 95% confidence intervals of odds ratios and multiplicity-corrected p-values from Fisher exact tests are provided in Table 1.

Figure 3

a.


TRASTUZUMAB
GSE50948
TRASTUZUMAB
GSE37946
TRASTUZUMAB
GSE42822
TRASTUZUMAB
TransNEO
VELIPARIB
GSE194040
VELIPARIB
GSE164458
OLAPARIB
GSE173839
0.06*
0.17
0.02**
0.06*
0.04**
0.09*
< 0.01**
Patient's predicted response scores
2.0
1.5
1.0
0.5
0.0
Non-Responder
Responder
Outcome
Outcome
Non-Responder
Responder

b.

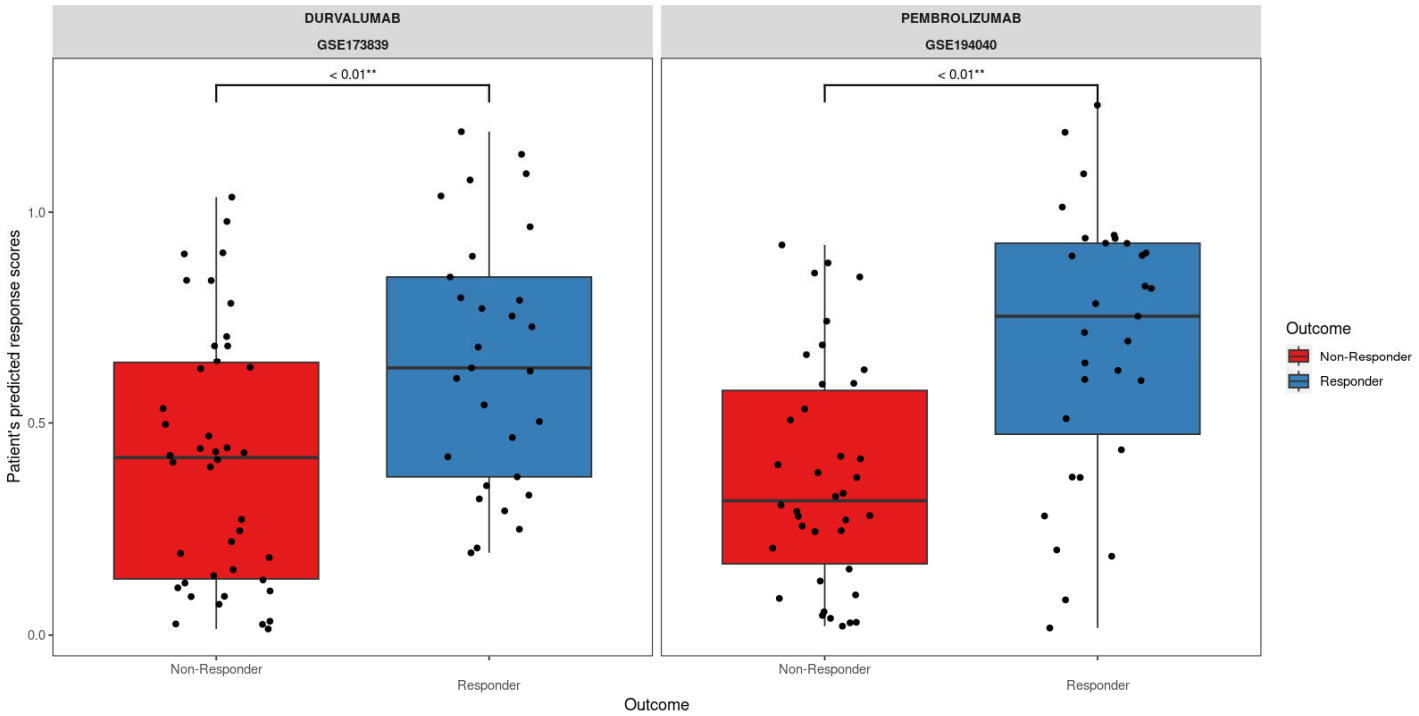


Figure 3: Patient's final predicted response scores of BC-SELECT among responders and non-responders. Red = non-responders, Blue = responders. **a. Targeted therapies.** X-axis shows the specific clinical trial, and Y-axis shows the distribution of predicted response scores within the trial. Final predicted BC-SELECT scores are significantly higher in responders compared to non-responders in 6 out of 7 trials. **b. Immunotherapy.** Final predicted BC-SELECT scores are significantly higher in responders compared to non-responders in both immunotherapy trials. Significance was determined via the Wilcoxon rank-sum test with correction for multiple hypothesis testing. Adjusted p-values are displayed in each box plot, where $p < 0.1$ is indicated by a single asterisk (*) and $p < 0.05$ by a double asterisk (**).

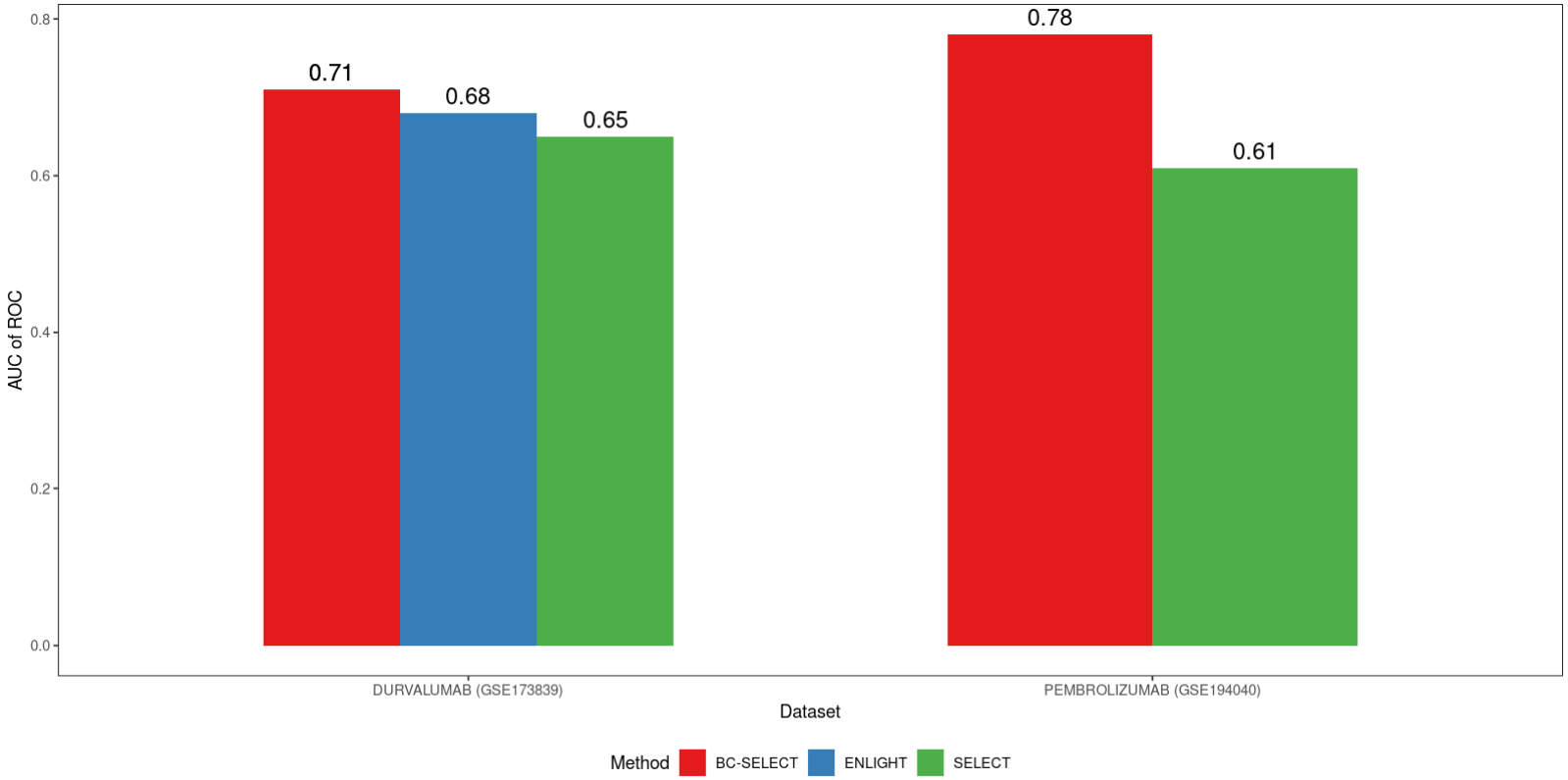


Figure 4: Performance comparisons between BC-SELECT and other algorithms in the context of immunotherapy. Red = BC-SELECT, blue = ENLIGHT, and green = SELECT. X-axis shows the specific clinical trial, and Y-axis shows the ROC-AUC. SELECT could be compared to BC-SELECT with both trials. ENLIGHT could be compared via durvalumab as this was a trial used by the developers for their validation. BC-SELECT outperforms SELECT/ENLIGHT in ROC-AUC values for both validation immunotherapy trials.

	DRUG	TARGET	TRIAL	GENETIC INTERACTION	Odds ratio	95% Confidence Interval	Fisher exact test adjusted P-value
Immuno therapy	DURVALUMAB	PDCD1, CD274	GSE173839[75] (R:29, NR:42)	SR	**3.273**	**(1.22, 8.782)**	**0.03**
	PEMBROLIZUMAB		GSE194040[13] (R:31, NR:38)	SR	**7.057**	**(2.428, 20.514)**	**< 0.01**
Targeted therapy	VELIPARIB	PARP1, PARP2	GSE194040[13] (R: 27, NR: 44)	SDL	**3.117**	**(1.151, 8.438)**	**0.07**
			GSE164458[74] (R:127, NR:110)	SL	**1.803**	**(1.049, 3.098)**	**0.07**
	OLAPARIB	PARP1, PARP2, PARP3	GSE173839[75] (R:29, NR:42)	SL	**3.542**	**(1.307, 9.6)**	**0.05**
	TRASTUZUMAB	ERBB2 (HER2)	GSE50948[70] (R:31, NR:32)	SDL	2.318	(0.839, 6.404)	0.18
			GSE37946[71] (R:27, NR:23)	SDL	2.263	(0.727, 7.047)	0.28
			GSE42822[72] (R:12, NR:13)	SDL	**11.25**	**(1.647, 76.852)**	**0.05**
			TransNEO[73] (R:19, NR:46)	SL	1.895	(0.643, 5.589)	0.28

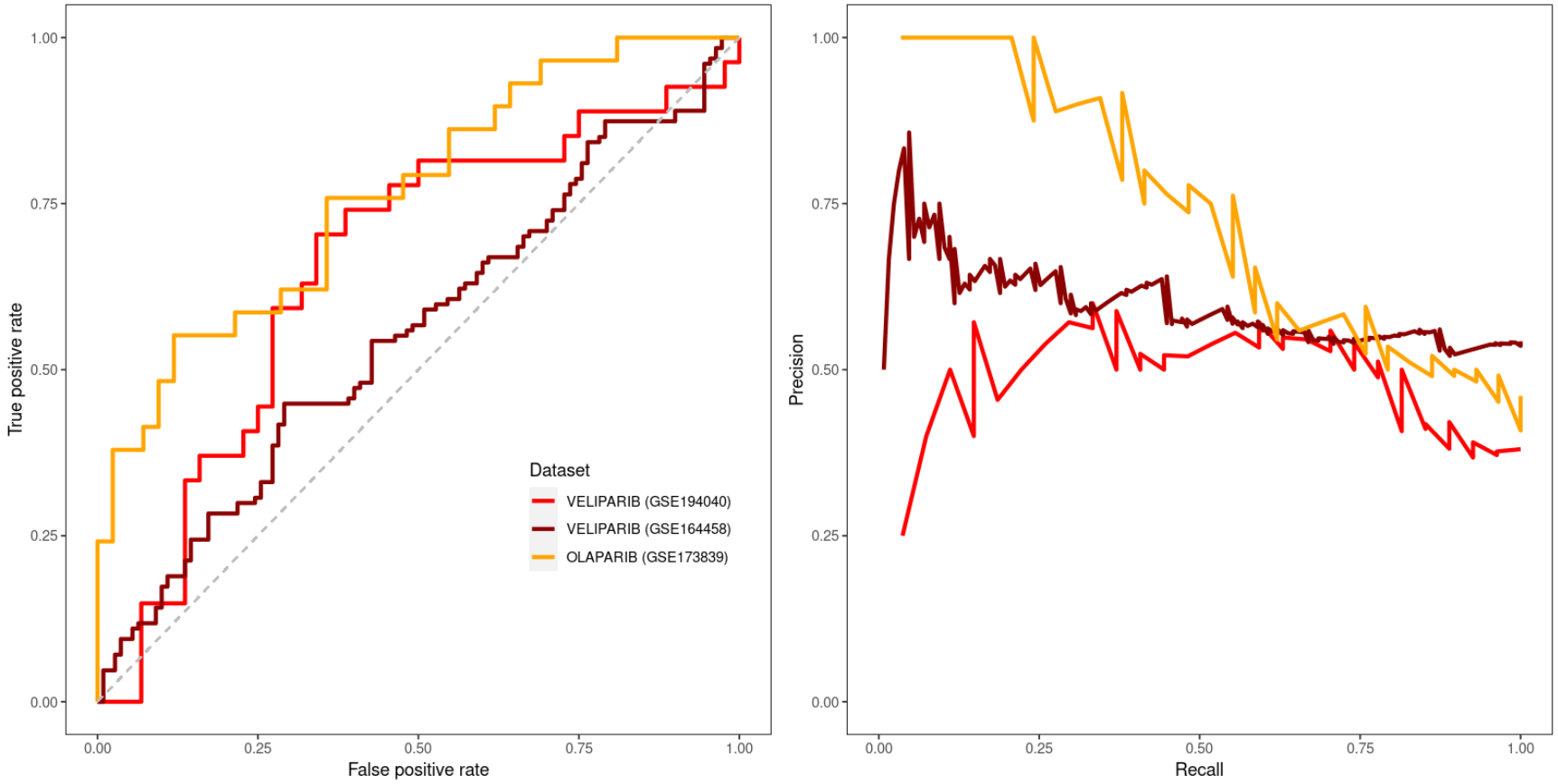
Table 1: Odds ratios between responders and non-responders using patients' final predicted response scores from BC-SELECT. R: Responders. NR: Non-Responders. SL: Synthetic Lethality. SDL: Synthetic Dosage Lethality. SR: Synthetic Rescue. All odds ratios were greater than 1, and six trials showed statistical significance (bold) based on 95% confidence interval. Last column presents the p-values of the Fisher exact test after multiple hypothesis adjustment. P-values meeting the 0.1 FDR threshold are highlighted in bold.

	DRUG	TARGET	TRIAL	GENETIC INTERACTION	BC-SELECT ROC-AUC (Empirical P-value)	Partner gene expression ratio ROC-AUC (Empirical P-value)	Target gene expression ROC-AUC (Empirical P-value)
Immuno therapy	DURVALUMAB	PDCD1, CD274	GSE173839[75] (R:29, NR:42)	SR	**0.71** (0.01)	**0.62** (0.01)	0.70 (0.03)
	PEMBROLIZUMAB		GSE194040[13] (R:31, NR:38)	SR	**0.78** (<0.01)	**0.66** (<0.01)	**0.75** (<0.01)
Targeted therapy	VELIPARIB	PARP1, PARP2	GSE194040[13] (R: 27, NR: 44)	SDL	**0.66** (0.02)	**0.59** (0.02)	0.63 (0.1)
			GSE164458[74] (R:127, NR:110)	SL	0.56 (0.32)	0.45 (0.99)	**0.61** (0.02)
	OLAPARIB	PARP1, PARP2, PARP3	GSE173839[75] (R:29, NR:42)	SL	**0.77** (<0.01)	**0.8** (<0.01)	0.56 (0.27)
	TRASTUZUMAB	ERBB2 (HER2)	GSE50948[70] (R:31, NR:32)	SDL	**0.63** (0.02)	0.6 (0.05)	0.6 (0.06)
			GSE37946[71] (R:27, NR:23)	SDL	0.59 (0.08)	0.48 (0.25)	0.6 (0.09)
			GSE42822[72] (R:12, NR:13)	SDL	**0.8** (0.01)	0.64 (0.38)	**0.78** (0.01)
			TransNEO[73] (R:19, NR:46)	SL	0.64 (0.19)	0.51 (0.79)	0.69 (0.04)

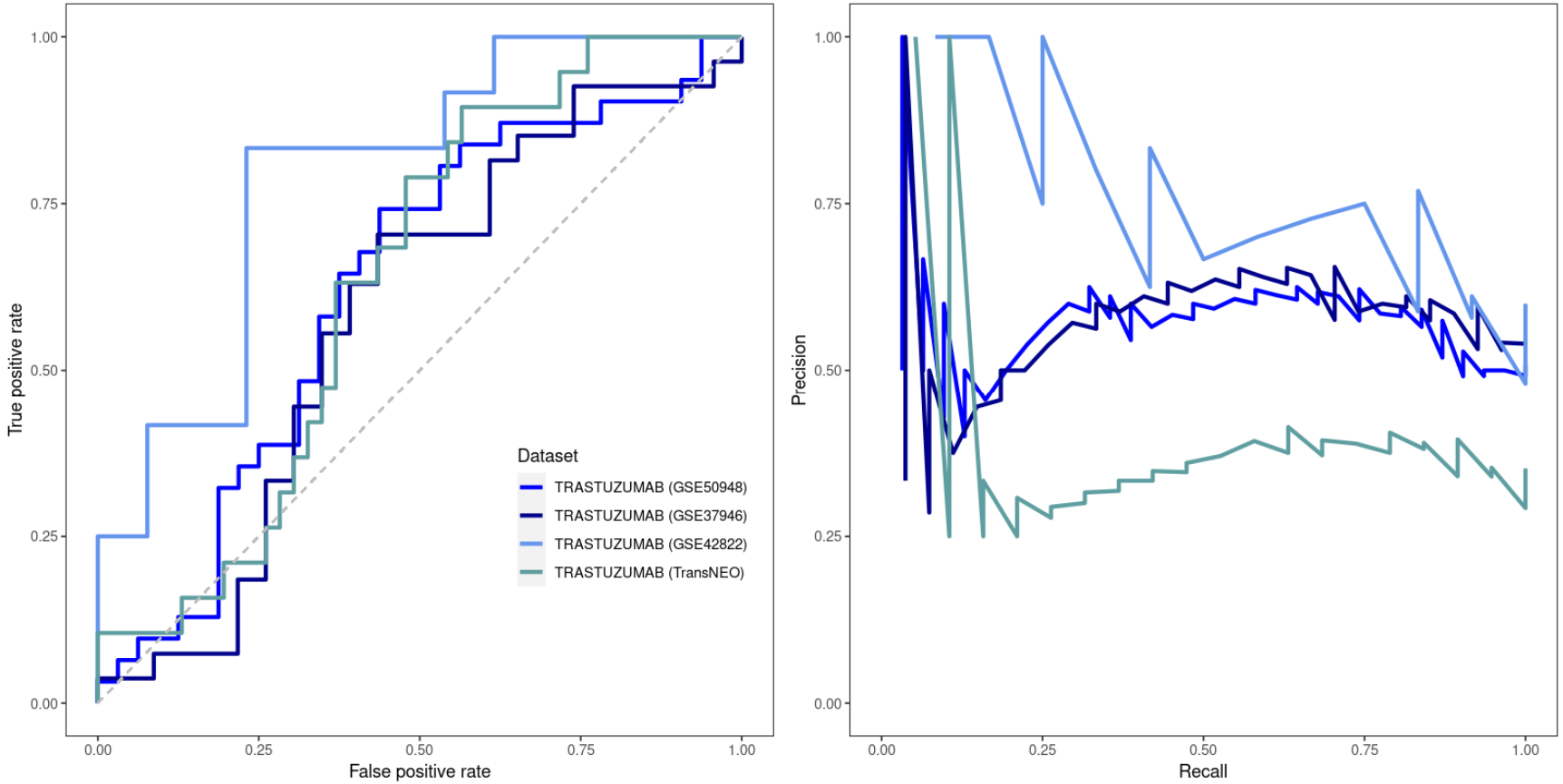
Table 2: ROC-AUC values and empirical p-values for BC-SELECT and its score components (partner gene expression ratio and target gene expression) in predicting nine breast cancer clinical trials
R: Responders. NR: Non-Responders. SL: Synthetic Lethality. SDL: Synthetic Dosage Lethality. SR: Synthetic Rescue. We report ROC-AUCs and empirical p-values, which describe the performance within a Monte Carlo-generated null distribution. ROC-AUC values are highlighted in bold if the ROC-AUC values fall within the top 2.5% of null AUC distributions. For targeted therapies, the reported ROC-AUC value is the higher one that was generated using either SL and SDL.

Supplementary Figure 1

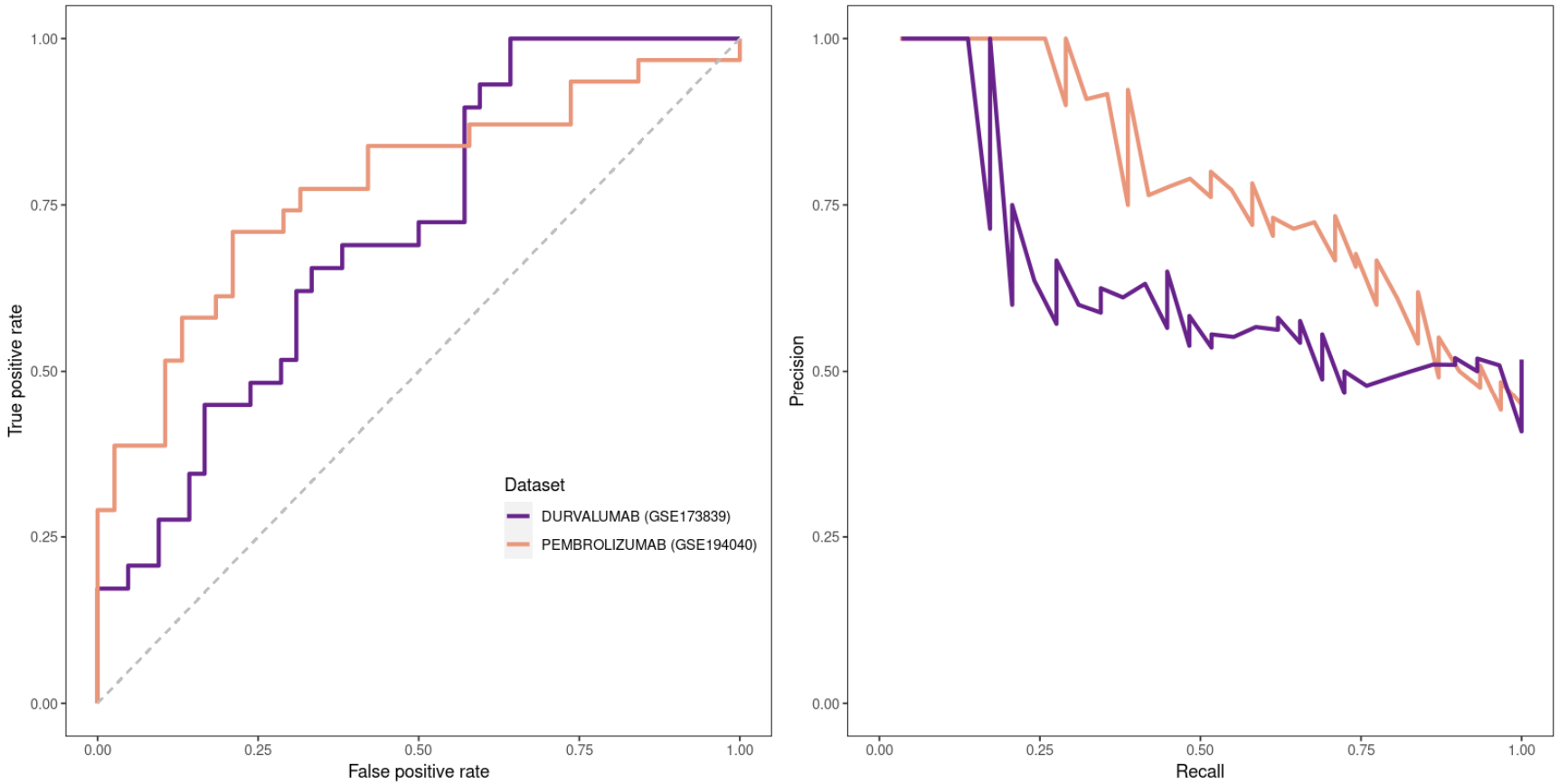
a.



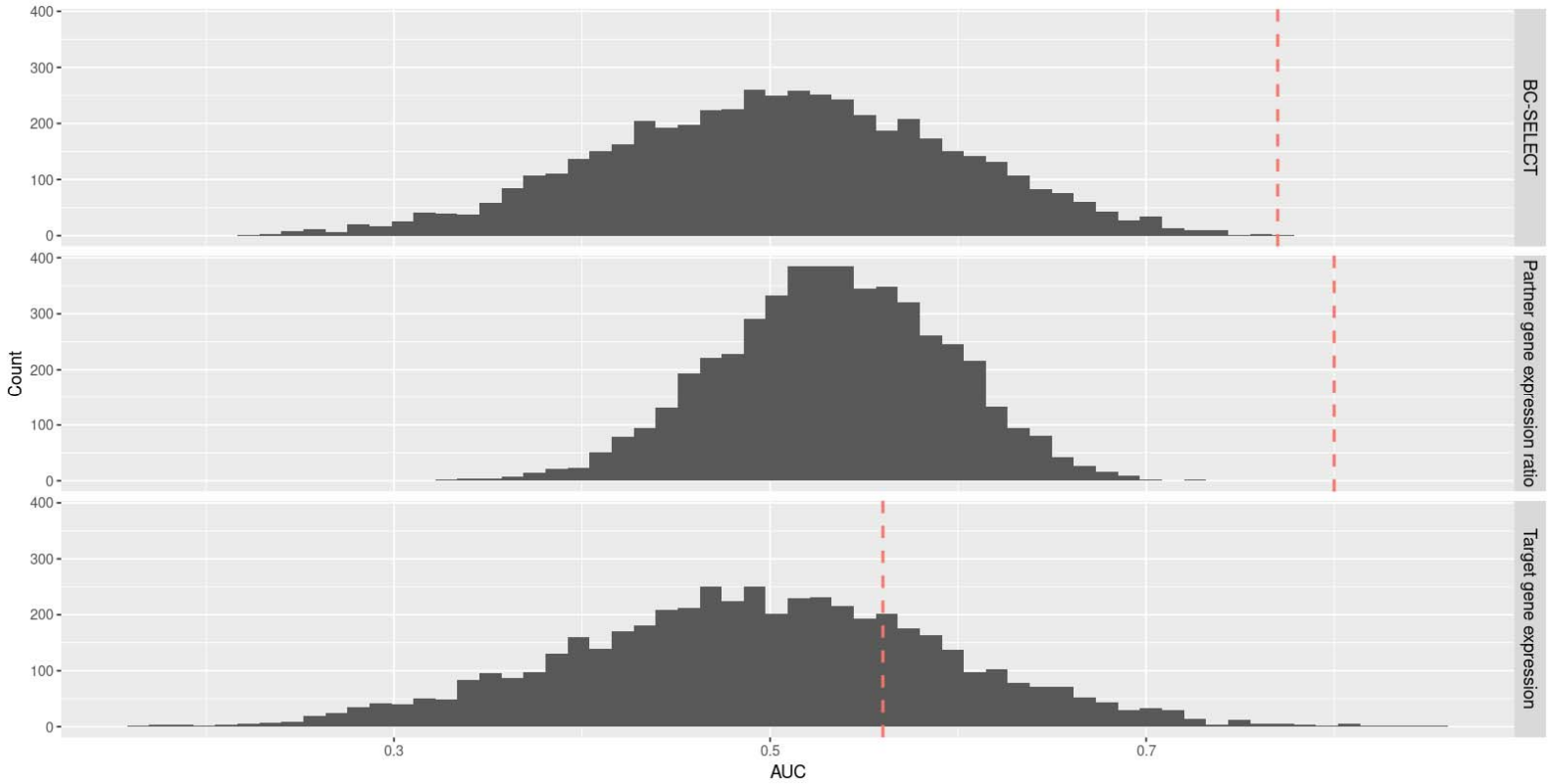
b.


True positive rate
False positive rate
0.00
0.25
0.50
0.75
1.00
Dataset
TRASTUZUMAB (GSE50948)
TRASTUZUMAB (GSE37946)
TRASTUZUMAB (GSE42822)
TRASTUZUMAB (TransNEO)
Precision
Recall

c.



Supplementary Figure 1: ROC and precision-recall curves across therapy types. The left panel shows the false positive rate plotted against the true positive rate for the ROC curve. The right panel shows the precision recall curve. **a. PARP inhibitors.** Veliparib-based trials are in red and olaparib is in orange. **b. HER2-directed therapy.** Trastuzumab-based trials are in blue. **c. Immunotherapy.** Durvalumab is represented in purple, and pembrolizumab is shown in pink.



Supplementary Figure 2: Example of Monte Carlo-based "null distributions" for assessment of BC-SELECT performance of olaparib (GSE173839[75]). The figure shows the distribution of 5000 ROC-AUCs (with AUC prediction on the X-axis and number of iterations with randomly selected genes for prediction achieving that AUC on the Y-axis). The null AUC distributions are centered around 0.5, indicative of random chance classification, while their varying right-tail thickness and the central tendency reflect differences associated with the efficacy of prediction using randomly selected genes for each variable. The three red dashed lines indicate the ROC-AUC values obtained by BC-SELECT, partner gene expression ratio, and target gene expression. Empirical p-values are calculated based on the proportion of the right tail area of the red dashed line. In this example, both BC-SELECT and partner gene expression ratio exhibit smaller empirical p-values and superior performance of synthetic lethality partner genes in drug response prediction compared to chance.