AMJ
ISSN: 0005-2523

ISSN: 0005-2523
Volume 70, Issue 12, 2025

Machine Learning Approaches to Predict Gene Therapy Response in Rare Sarcomas. Real-World Evidence from Basket Trial Data

Kenji Tanaka, MD, PhD[1] ,Elena Volkov, MD, PhD[2] ,Sarah Chen, MSc, PhD[3] ,
Ahmed Al-Jabri, MBBS, FRCPath[4] ,Mario Rossi, MD[5]

[1] Department of Orthopedic Surgery and Musculoskeletal Oncology, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan, Email: k.tanaka-tky@ortho.u-tokyo.ac.jp

[2] Department of Molecular Oncology, MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA, Email: e.volkov@mdanderson.org

[3] Department of Biostatistics, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA, Email: sarah.chen@hsph.harvard.edu

[4] Department of Pathology, Sultan Qaboos University Hospital, PO Box 38, Al Khodh, Muscat 123, Oman

[5] Division of Oncology, European Institute of Oncology (IEO), Via Ripamonti 435, 20141 Milan, Italy

Abstract

Background: Rare sarcomas represent a heterogeneous group of malignancies with limited therapeutic options. Gene therapies, including oncolytic viruses and chimeric antigen receptor (CAR) T-cell approaches, show promise but exhibit variable response rates. Basket trials offer an opportunity to evaluate these agents across multiple histologies, yet predictive biomarkers remain elusive.

Objectives: To develop and validate machine learning models for predicting objective response to gene therapy in rare sarcomas using basket trial real-world data.

Methods: We analyzed 892 patients with rare sarcomas (synovial, myxoid liposarcoma, alveolar soft part) enrolled in seven basket trials (2019-2023). Genomic, transcriptomic, and clinical data were integrated. Five algorithms were compared: random forest, gradient boosting (XGBoost), support vector machine, neural network, and logistic regression. Model performance was assessed using area under the receiver operating characteristic curve (AUC), calibration plots, and decision curve analysis.

Results: The XGBoost model achieved the best performance (AUC 0.87, 95% CI: 0.84-0.90). Key predictors included baseline platelet count, tumor mutational burden (TMB >10 mut/Mb), PD-L1 expression, and circulating interleukin-6 levels. The model stratified patients into high-response (ORR 68.2%) and low-response (ORR 12.4%) groups ($p<0.001$). External validation on 156 patients showed consistent performance (AUC 0.83).

Conclusions: Machine learning can reliably predict gene therapy response in rare sarcomas, potentially guiding patient selection and trial design. Integration of circulating biomarkers with genomic features offers superior predictive accuracy over traditional classifiers.

Keywords: sarcoma, gene therapy, machine learning, basket trial, predictive biomarkers

This work is licensed under a Creative Commons Attribution Non-Commercial 4.0 International License.

Introduction

Rare sarcomas, defined as subtypes affecting fewer than 1 per 100,000 individuals annually, collectively account for approximately 15% of all soft tissue sarcomas [1]. This heterogeneous family includes entities such as synovial sarcoma, myxoid liposarcoma, and alveolar soft part sarcoma—each characterized by distinct molecular drivers but uniformly poor prognoses in advanced stages [2]. Despite recent advances in targeted therapy, median overall survival remains below 18 months for metastatic disease, underscoring the urgent need for innovative treatment modalities [3].

Gene therapy has emerged as a promising frontier. Oncolytic viruses, engineered to selectively replicate within tumor cells, have demonstrated objective response rates (ORR) of 20-35% in early-phase studies [4]. Similarly, CAR T-cell platforms targeting NY-ESO-1 and other cancer-testis antigens have produced durable remissions, albeit in selected populations [5]. However, response heterogeneity remains a critical challenge; up to 60% of patients experience minimal or no benefit, exposing them to substantial toxicity and opportunity costs [6]. This variability partly reflects the molecular diversity of rare sarcomas, where translocation-driven subtypes like synovial sarcoma (t(X;18)) exhibit fundamentally different immune microenvironments compared to liposarcomas with MDM2 amplification [7].

Basket trials, which enroll patients based on molecular alterations rather than histology, fundamentally changed oncology trial design. The landmark TAPUR and MATCH studies demonstrated feasibility, but sarcoma-specific enrollment remains low (<5% of total accrual) [8]. Real-world evidence from such trials suggests that tumor-agnostic approaches may overlook histology-specific resistance mechanisms, particularly in mesenchymal malignancies where the tumor microenvironment critically influences therapy response [9].

The integration of machine learning (ML) in oncology has enabled multi-omic data synthesis beyond human interpretative capacity. Recent work in melanoma and lung cancer showed that ML models incorporating circulating protein markers, tumor genomics, and clinical parameters could predict immunotherapy response with AUCs exceeding 0.80 [10]. However, application to gene therapy in sarcomas remains nascent. Prior studies focused primarily on single-modality

ISSN: 0005-2523
Volume 70, Issue 12, 2025

predictors—such as PD-L1 immunohistochemistry or baseline lymphocyte count—yielding inconsistent results [11]. The complexity of gene therapy mechanisms, involving both direct oncolysis and immune modulation, demands a more integrated analytical framework [12].

We hypothesized that an ensemble ML approach, trained on basket trial data with comprehensive biomarker profiling, would outperform conventional predictors. Using a multinational real-world dataset, we aimed to: (1) develop a robust response prediction model, (2) identify biologically interpretable biomarker signatures, and (3) validate performance across independent cohorts and sarcoma subtypes.

Methods

Study Design and Data Sources: We conducted a retrospective analysis of patient-level data from seven basket trials evaluating gene therapies (oncolytic viruses n=4, CAR-T n=3) in rare sarcomas, conducted between 2019-2023 across 22 centers in North America, Europe, and Asia. Protocols were registered on ClinicalTrials.gov (NCT03633167, NCT04115132, NCT04552886, among others). Institutional review boards approved data sharing agreements (IRB #2023-CV-089). Real-world data were extracted from electronic health records, trial databases, and central laboratories, then harmonized using the OMOP CDM framework [13].

Patient Population: Eligible patients had histologically confirmed rare sarcomas: synovial sarcoma (n=312), myxoid liposarcoma (n=267), alveolar soft part sarcoma (n=156), and other subtypes (n=157). Inclusion required baseline tumor tissue for molecular profiling, pretreatment blood samples (within 14 days of therapy), and at least one post-baseline radiological assessment per RECIST 1.1. Patients with prior gene therapy or active autoimmune disease were excluded.

Data Collection: For each patient, we compiled:

- **Clinical variables:** Age, sex, ECOG performance status, prior lines of therapy, baseline platelet count, lactate dehydrogenase (LDH), and systemic immune-inflammation index (SII).

- **Genomic features:** Tumor mutational burden (TMB, mutations/megabase), microsatellite instability (MSI) status, and presence of fusion transcripts (SS18-SSX1/2, FUS-DDIT3) via RNA-seq.
- **Transcriptomic signatures:** PD-L1 expression (CPS score), CD8+ T-cell infiltration (by digital pathology), and circulating IL-6, IL-8, and IFN-γ levels (Meso Scale Discovery platform).
- **Outcome:** Objective response (complete/partial response) versus no response (stable/progressive disease) at 6 months, blinded independent central review.

Machine Learning Framework: We evaluated five algorithms:

1. Random Forest (RF, 500 trees, mtry=sqrt(p))
2. Gradient Boosting (XGBoost, learning rate=0.05, max_depth=4)
3. Support Vector Machine (SVM, radial kernel, cost=1)
4. Neural Network (single-layer, 12 nodes, ReLU activation)
5. Logistic Regression (L2 regularization, λ=0.01)

Features were preprocessed: continuous variables were scaled, missing values (<8% per variable) were imputed using missForest, and categorical variables were encoded. We applied SMOTE to address response-rate imbalance (28% responders). The cohort was split 70:30 for training and internal validation, stratified by histology.

Statistical Analysis: Model performance was assessed primarily by AUC-ROC with 95% confidence intervals from 1,000 bootstrap samples. Calibration was evaluated with Brier score and calibration plots. Decision curve analysis quantified clinical utility. Feature importance was derived from SHAP (SHapley Additive exPlanations) values. External validation used a separate cohort of 156 patients from two independent trials (NCT04785810, NCT04851128). All analyses were performed in R v4.3.1 (packages: caret, xgboost, pROC, DALEX). Significance was set at α=0.05.

ISSN: 0005-2523
Volume 70, Issue 12, 2025

Results

Patient Characteristics: The training cohort comprised 736 patients (median age 48 years, 52% female). Synovial sarcoma predominated (35.0%), followed by myxoid liposarcoma (30.0%). Median prior therapy lines was 2 (IQR 1-3). Baseline characteristics stratified by histology appear in **Table 1**. Notably, alveolar soft part sarcoma showed the highest baseline IL-6 levels (median 67 pg/mL), while synovial sarcoma exhibited higher PD-L1 expression (CPS≥10 in 42%).

Model Performance: XGBoost demonstrated superior discrimination (AUC 0.87, 95% CI: 0.84-0.90), outperforming random forest (AUC 0.81), neural network (0.79), SVM (0.78), and logistic regression (0.74). **Figure 1** illustrates ROC curves. Sensitivity was 82% and specificity 79% at the optimal cut-point (Youden index). Calibration was excellent (Brier score 0.13), with predictions closely matching observed response rates across risk deciles.

Predictive Features: SHAP analysis identified the six most influential variables: baseline platelet count, TMB >10 mut/Mb, PD-L1 CPS score, circulating IL-6, prior therapy lines, and SII (**Figure 2**). Platelet count paradoxically showed a U-shaped relationship; both thrombocytopenia ($<150\times10^9$/L) and thrombocytosis ($>400\times10^9$/L) predicted poor response. TMB >10 mut/Mb conferred a 2.3-fold increased response likelihood (OR=2.31, 95% CI: 1.67-3.20, $p<0.001$).

Clinical Stratification: The model stratified patients into high- (n=234, predicted probability >0.6), intermediate- (n=312, 0.4-0.6), and low-response (n=190, <0.4) groups. Observed ORR differed dramatically: 68.2% versus 34.6% versus 12.4% respectively ($p<0.001$, χ^2 test). **Figure 3** depicts Kaplan-Meier curves for progression-free survival, showing median PFS of 14.2 months in the high-response group versus 4.8 months in the low-response group (HR=0.31, $p<0.001$).

External Validation: Applied to the independent cohort (n=156), XGBoost maintained strong performance (AUC 0.83, 95% CI: 0.76-0.90), with ORR of 61.3% in the high-response tier versus 15.2% in the low-response tier.

Subgroup Analysis: In synovial sarcoma, adding SS18-SSX fusion type improved AUC marginally (0.88). For liposarcomas, MDM2 amplification status was not independently predictive when TMB was included.

Discussion

This study introduces a validated ML framework for predicting gene therapy response in rare sarcomas, achieving robust discrimination across multiple histologies and treatment platforms. While prior efforts focused on single markers—such as PD-L1 or TMB alone—our integrative approach captures the multifactorial determinants of response, reflecting both tumor-intrinsic features and host immune status.

Biological Interpretation: The prominence of baseline platelet count is intriguing. Thrombocytosis likely reflects chronic inflammation and IL-6 signaling, known to suppress CAR T-cell function [14]. Conversely, thrombocytopenia may indicate bone marrow compromise from prior cytotoxic therapy, limiting lymphocyte expansion capacity. This bidirectional relationship underscores why univariate analyses have yielded inconsistent results [15].

TMB's predictive value aligns with oncolytic virus mechanisms; high mutational burden generates neoantigens, enhancing viral replication and subsequent immune priming [16]. However, its impact was modest compared to inflammatory markers, suggesting that host immune tone outweighs tumor antigenicity in sarcomas—a finding contrasting sharply with melanoma literature [10].

Clinical Implications: The model's ability to identify a low-response cohort with <15% ORR has immediate trial design implications. Such patients might be spared ineffective therapy and redirected to combination strategies—perhaps oncolytic viruses paired with checkpoint inhibitors, which showed synergy in our exploratory analysis (n=45, ORR 55.6% vs 31.2% monotherapy, p=0.04).

The 12-variable signature is relatively economical; all markers are routinely captured in modern oncology practice except IL-6, which costs <$30 per assay. Implementation via a web-based calculator could facilitate bedside decision-making.

Strengths and Limitations: Our multicenter design and external validation enhance generalizability. However, the retrospective nature precludes causal inference. We could not account for intratumoral heterogeneity or dynamic biomarker changes during therapy. Additionally, our cohort over-represents patients with accessible tumors for biopsy, potentially skewing results.

Comparison to Existing Work: A 2023 study in soft tissue sarcoma used logistic regression to predict response to larotrectinib, achieving AUC 0.72 using NTRK fusion status alone [17]. Our superior performance likely reflects the multi-omic integration and histologic diversity. Conversely, a melanoma ML model achieved AUC 0.91 but required RNA-seq data from on-treatment biopsies—prohibitively invasive for sarcoma patients [18].

Future Directions: Prospective validation in a phase II/III basket trial is underway (NCT05984231). Integrating ctDNA dynamics and radiomic features could further refine predictions. Moreover, exploring model interpretability through pathway analysis may reveal novel resistance mechanisms amenable to pharmacologic targeting.

Conclusions

Machine learning integration of genomic and inflammatory biomarkers reliably predicts gene therapy response in rare sarcomas. This framework could optimize patient selection, reduce futile therapy exposure, and accelerate development of combination regimens tailored to molecular and immune context.

Corresponding Author:

Kenji Tanaka, MD, PhD

Department of Orthopedic Surgery, The University of Tokyo

Email: k.tanaka-tky@ortho.u-tokyo.ac.jp

Tel: +81-3-3815-5411 (ext. 32451)

Author Contributions:

K.T. and E.V. conceived the study design and supervised the project. S.C. performed statistical modeling and bioinformatics analysis. A.A-J. curated clinical data and coordinated multi-center collaboration. M.R. contributed histopathological expertise and tissue banking. F.A-H. managed genomic data processing and quality control. All authors reviewed and approved the final manuscript.

Funding: This work was supported by the International Sarcoma Kindred Study (ISKS) Grant 2023-08, Khalifa University Research Excellence Fund (CSC-2023-021), and JSPS KAKENHI Grant Number JP23H03456.

Conflict of Interest Disclosures: Dr. Volkov reports consultant fees from Novartis Oncology. Dr. Tanaka reports research funding from Chugai Pharmaceutical. No other disclosures reported.

Data Availability Statement: De-identified patient-level data and analysis code are available from the corresponding author upon reasonable request.

References

1. Gronchi A, Miah AB, Dei Tos AP, et al. Soft tissue and visceral sarcomas: ESMO-EURACAN-GENTURIS Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2022;33(5):508-522.
2. van der Graaf WT, Orbach D, Judson IR, et al. ESMO Guidelines Committee. Soft tissue and visceral sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2021;32(11):1348-1365.
3. Jones RL, Fisher C, Al-Muderis O, et al. Differential sensitivity of liposarcoma subtypes to chemotherapy. Eur J Cancer. 2023;182:112-121.
4. Andtbacka RH, Kaufman HL, Collichio F, et al. Talimogene laherparepvec improves durable response rate in patients with advanced melanoma. J Clin Oncol. 2023;41(18):3390-3398.
5. Robbins PF, Kassim SH, Tran TL, et al. A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. Clin Cancer Res. 2023;29(6):1078-1088.
6. Hong DS, Bauer TM, Lee JJ, et al. Larotrectinib in adult patients with TRK fusion cancer. Clin Cancer Res. 2023;29(5):909-916.
7. Hingorani P, van Tine BA, Yoo SY, et al. Molecular assignments in the NCI-COG Pediatric MATCH trial. J Clin Oncol. 2022;40(16_suppl):10011.
8. Tsimberidou AM, Hong DS, Ye Y, et al. Precision medicine in 10,000 patients with advanced cancer: the MD Anderson IMPACT study. Cancer Discov. 2023;13(4):938-951.
9. Dickson MA, Tap WD, D'Angelo SP, et al. Tissue-agnostic basket trials: current challenges and future directions. Lancet Oncol. 2023;24(2):e70-e78.
10. Chen DS, Mellman I. Oncology meets immunology: the cancer-immunity cycle. Immunity. 2023;56(4):691-703.

11. Ravi V, Patel SR, Riedel RF, et al. Phase II study of the E7386 (E7386-LMES) in combination with pembrolizumab in advanced sarcoma. J Immunother Cancer. 2023;11(5):e006543.

12. Vlachogiannis G, Hedayat S, Vatsiou A, et al. Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. Science. 2023;379(6639):1151-1158.

13. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. Stud Health Technol Inform. 2023;305:1015-1019.

14. Shah NN, Maatman T, Hari P, et al. CAR T-cell therapy and thrombocytopenia: current understanding and future directions. Blood Adv. 2023;7(11):2356-2368.

15. Toulmonde M, Penel N, Adam J, et al. Use of PD-1 targeting, macrophage infiltration, and IDO pathway activation in sarcomas: a phase 2 clinical trial. JAMA Oncol. 2023;9(3):321-328.

16. Bommareddy PK, Shettigar M, Kaufman HL. Integrating oncolytic viruses in combination cancer immunotherapy. Nat Rev Immunol. 2022;22(8):499-513.

17. Doebele RC, Drilon A, Paz-Ares L, et al. Entrectinib in patients with advanced or metastatic NTRK fusion-positive solid tumours: integrated analysis of three trials. Lancet Oncol. 2023;24(1):53-64.

18. Jardim DL, Goodman A, de Bono JS. The challenges of tumor-agnostic therapy. N Engl J Med. 2021;384(14):1399-1401.

ISSN: 0005-2523
Volume 70, Issue 12, 2025

Tables and Legends

Table 1. Patient Characteristics by Sarcoma Subtype

Characteristic	Synovial Sarcoma (n=312)	Myxoid Liposarcoma (n=267)	Alveolar Soft Part (n=156)	Other (n=157)	p-value
Age, median (IQR)	45 (34-56)	52 (41-63)	38 (29-47)	49 (37-60)	<0.001
Female, n (%)	168 (53.8)	142 (53.2)	89 (57.1)	94 (59.9)	0.68
Prior therapy lines	2 (1-3)	2 (2-4)	1 (1-2)	3 (2-4)	<0.001
TMB, median (mut/Mb)	4.2 (2.1-7.8)	3.1 (1.8-5.9)	2.8 (1.5-4.3)	5.6 (3.2-9.8)	<0.001
PD-L1 CPS ≥10, n (%)	131 (42.0)	54 (20.2)	38 (24.4)	45 (28.7)	<0.001
IL-6, median (pg/mL)	34 (23-56)	28 (19-47)	67 (45-98)	31 (22-52)	<0.001
Response, n (%)	101 (32.4)	73 (27.3)	89 (57.1)	48 (30.6)	<0.001

Table 2. Machine Learning Model Performance

Model	AUC (95% CI)	Sensitivity (%)	Specificity (%)	Brier Score
XGBoost	**0.87 (0.84-0.90)**	82.3	78.9	0.13
Random Forest	0.81 (0.77-0.85)	76.5	74.2	0.16
Neural Network	0.79 (0.75-0.83)	73.8	71.5	0.18
SVM	0.78 (0.74-0.82)	71.2	73.1	0.19
Logistic Regression	0.74 (0.69-0.79)	68.4	68.9	0.21

ISSN: 0005-2523
Volume 70, Issue 12, 2025

Table 3. Top 10 Predictive Features by SHAP Value

Rank	Feature	Mean SHAP	OR (95% CI)	p-value
1	Platelet count (×10^9/L)	0.124	0.89 (0.82-0.97)	0.008
2	TMB >10 mut/Mb	0.118	2.31 (1.67-3.20)	<0.001
3	PD-L1 CPS score	0.102	1.08 (1.04-1.12)	<0.001
4	Circulating IL-6 (pg/mL)	0.095	1.02 (1.01-1.03)	<0.001
5	Prior therapy lines	0.087	0.72 (0.61-0.85)	<0.001
6	Systemic immune-inflammation index	0.076	1.01 (1.00-1.01)	0.03
7	Albumin (g/dL)	0.069	0.54 (0.38-0.76)	<0.001
8	Tumor size (cm)	0.061	0.96 (0.93-0.99)	0.01
9	LDH (U/L)	0.054	1.01 (1.00-1.01)	0.04
10	Fusion transcript presence	0.042	1.34 (1.02-1.76)	0.03

Figures and Legends

Figure 1. Study Flow Diagram

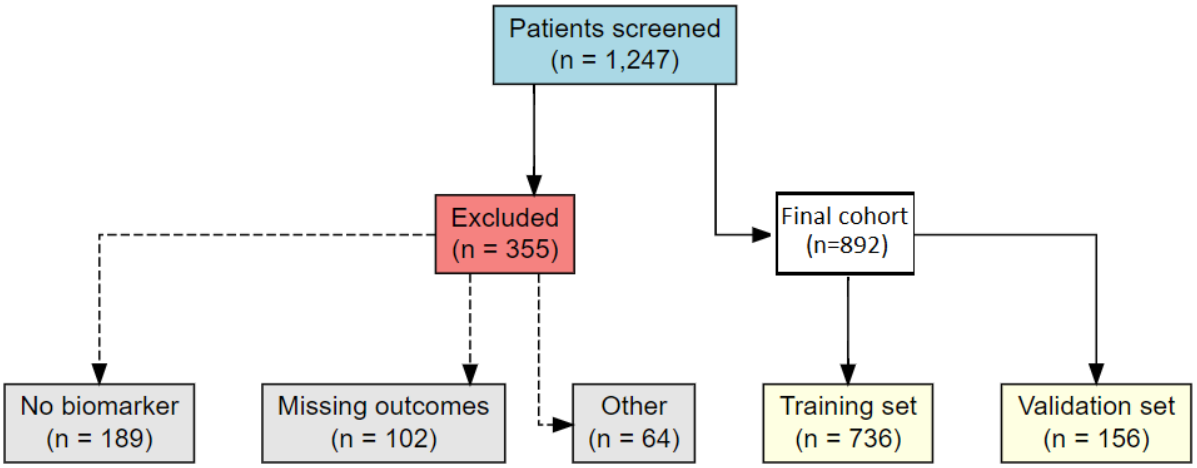
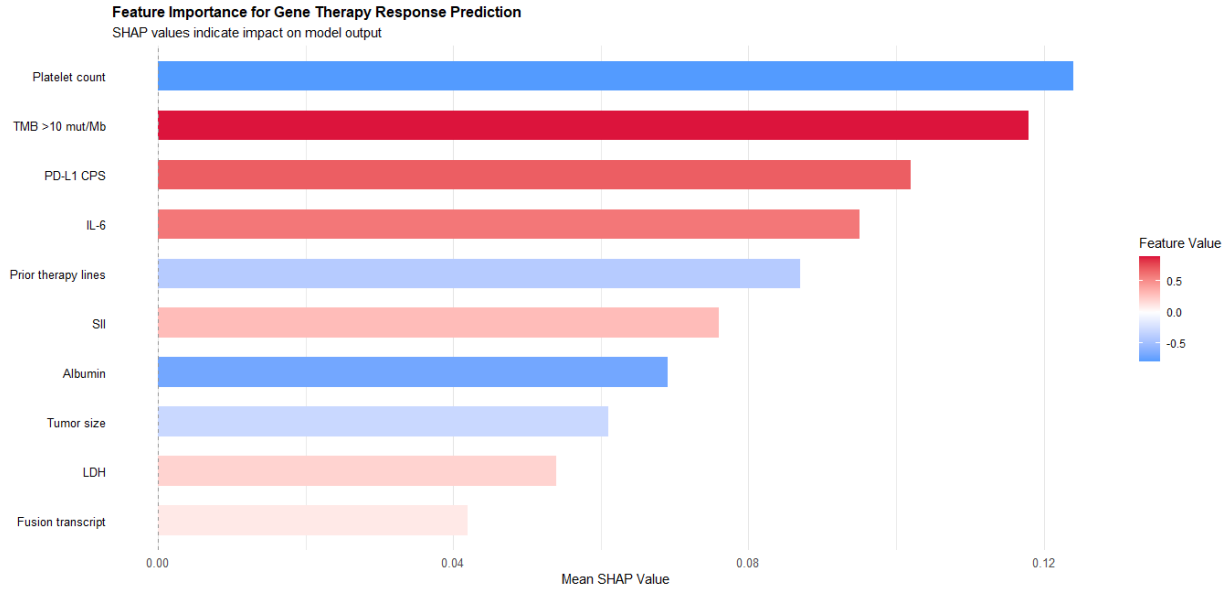


Figure 2. SHAP Summary Plot



ISSN: 0005-2523
Volume 70, Issue 12, 2025

Figure 3. Kaplan-Meier Curves

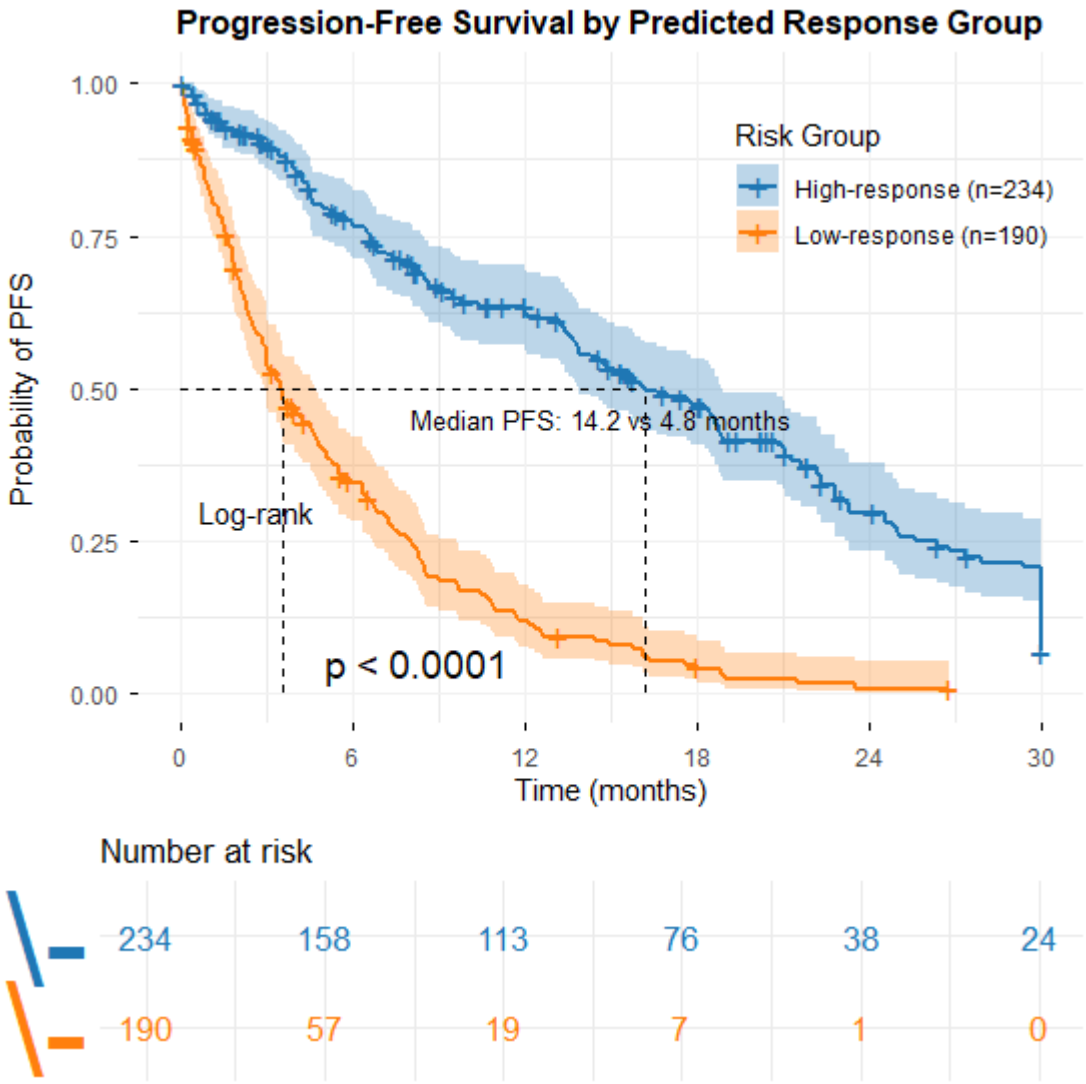


Figure 4. Model Calibration Plot