The Breast 86 (2026) 104715

Contents lists available at ScienceDirect

The Breast

journal homepage: www.journals.elsevier.com/the-breast

Predicting progression-free survival in hormone-receptor positive (HR+/HER2–) metastatic breast cancer (MBC) treated with CDK4/6 inhibitors: A machine learning approach

Sergio Pannunzio [a], Luca Mastrantoni [b,*], Noemi Maliziola [b], Letizia Pontolillo [b], Giovanna Garufi [b], Elena Di Monte [b], Alessandra Emiliani [a], Margherita Sgambato [b], Anna Cardillo [a], Antonella Palazzo [b], Armando Orlandi [b], Giampaolo Tortora [b], Emilio Bria [a,b]

[a] Ospedale Isola Tiberina – Gemelli Isola, Roma, Italy
[b] Comprehensive Cancer Center, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Università Cattolica del Sacro Cuore, Roma, Italy

ARTICLE INFO

Keywords:
Hormone receptor–positive metastatic breast cancer
CDK4/6 inhibitors
Machine learning
CoxNet regression
Gradient boosting machine
Predictive modelling

ABSTRACT

Background: In HR+/HER2– metastatic breast cancer (MBC), CDK4/6 inhibitors combined with endocrine therapy (ET) significantly improve progression-free survival (PFS). Machine learning (ML) approaches may improve individualized progression risk estimation.

Methods: We retrospectively analysed HR+/HER2– MBC patients treated with first-line CDK4/6i plus ET to develop CoxNet regression and Gradient Boosting Machine (GBM) models from baseline clinicopathological features. The primary endpoint was PFS prediction. The dataset was split into a 70/30 train/validation set. Performance was assessed by Harrell's C-index (1000 bootstrap replicates). Risk stratification was performed using Gaussian Mixture Modeling (GMM) to define high- and low-risk groups. Cox regression estimated the corresponding hazard ratios (HR). Early progression at 6 months (EP) prediction was evaluated using the area under the receiver operating characteristic curve (AUROC).

Results: 459 patients were included, with a median follow-up of 43.7 months (95% CI 39.6–48.3). Median PFS was 29.3 months (95% CI 24.0–33.7). Both ML models achieved strong predictive performance, with a Harrell's C-index of 0.74 (95% CI 0.67–0.80) in the validation set. The main predictors were liver metastases, Ki67 expression, and primary endocrine resistance. Stratification defined two risk groups with significantly different PFS in the validation set (HR 2.58, 95% CI 1.65–4.03, $p = 3.3 \times 10^{-5}$). Median PFS was 34.8 (95%CI 24.0–52.4) in the low-risk and 10.6 (95%CI 7.7–14.6) in the high-risk group. For EP prediction, the model achieved an AUROC of 0.77 (95% CI 0.61–0.89).

Conclusions: This study supports the clinical applicability ML models using baseline clinicopathological variables for individualized risk stratification in HR+/HER2– MBC.

1. Introduction

Breast cancer (BC) remains the most commonly diagnosed malignancy and the leading cause of cancer-related death among women worldwide [1]. The hormone receptor-positive, HER2-negative (HR+/HER2−) subtype accounts for approximately 65% of all BC cases [2,3].

Pivotal phase III trials have consistently demonstrated the superiority of CDK4/6i combined with aromatase inhibitors (AIs) or fulvestrant in prolonging progression-free survival (PFS) compared with ET alone, with pooled hazard ratios (HR) of 0.55 and 0.58, respectively [4].

Despite the broad efficacy, approximately 10% of patients experience primary resistance, progressing within the first few months of treatment [5]. Resistance to CDK4/6i may be driven by multiple molecular alterations, including biallelic RB1 loss [6,7,8], CCNE1/2 overexpression [9,10,11], FAT1 deletions [12], and hyperactivation of the FGFR, PI3K/AKT, MYC amplifications and RAS signalling pathways [13,14]. While PI3K pathway activation is a key endocrine resistance mechanism, PIK3CA mutations have not been directly linked to CDK4/6i resistance [15]. Although ESR1 and TP53 mutations are frequent in HR

* Corresponding author. Fondazione Policlinico Universitario Agostino Gemelli, Largo Agostino Gemelli, n. 8, 00168, Rome, Italy.
E-mail address: luca.mastrantoni01@icatt.it (L. Mastrantoni).

https://doi.org/10.1016/j.breast.2026.104715
Received 13 November 2025; Received in revised form 11 January 2026; Accepted 26 January 2026
Available online 30 January 2026

+ disease, their predictive role in CDK4/6i resistance remains unclear [16]. Moreover basal-like tumors, even when HR+ (2.6% in the MONALEESA dataset), demonstrate limited response to these agents [17].

In this context, machine learning (ML) has gained growing interest as a strategy for integrating clinical, molecular, imaging, and pathological data to outperform prediction [18,19]. ML applications have demonstrated promising results in anticipating recurrence, toxicity, and neoadjuvant response [20,21,22,23]. Recent studies suggest that using an extreme gradient boosting (XGBoost) tree algorithm may even support de-escalation strategies in select patients early BC [24,25].

Developing accurate ML-based models may help identify high-risk patients and guide therapeutic decision-making in HR+/HER2− metastatic breast cancer.

2. Materials & methods

2.1. Patients' population

This is a retrospective cohort study aimed at developing prediction models in HR-positive/HER2-negative metastatic BC. The study has been conducted and reported in accordance with the guidelines for transparent reporting of multivariable prediction models for individual prognosis or diagnosis (TRIPOD-AI) [26].

Inclusion criteria. Patients who fulfilled the following inclusion criteria were included: (1) pathologically confirmed diagnosis of HR-positive/HER2-negative metastatic BC between April 2017 and December 2024; (2) patients receiving CDK4/6 inhibitors in association to AI or fulvestrant in first line metastatic BC. Information and survival data were retrospectively collected from electronic medical records; clinical data were collected until April 2017, and each patient's follow-up was updated until March 2025.

Patient's cohorts. Patients treated at two Italian institutions, Fondazione Policlinico Universitario A. Gemelli IRCCS and Unit of Ospedale Isola Tiberina-Gemelli Isola were considered. The collection of anonymous data was approved by the institutional ethical review board (Prot ID3315 n.0029524/20 July 15th, 2020).

2.2. Clinical data

Measurements. The following clinical variables were collected: menopausal status, age, hormone resistance (primary vs secondary), histology, estrogen receptor and progesterone receptor expression, HER2 status (HER2 0 vs low), Ki67 value, sites of metastasis, CDK4/6 inhibitor, hormonal therapy, synchronous or metachronous metastasis. Women were considered postmenopausal if they met one of the following requirements: (1) age less than 60 years and amenorrhea for at least 12 months in the absence of chemotherapy, tamoxifen or ovarian suppression; (2) previous bilateral oophorectomy; (3) age over 60 years. Hormone resistance for early breast cancer was defined as primary resistance for patients who relapse during the first 2 years of adjuvant ET; secondary (acquired) resistance was defined as a relapse that occurs after at least 2 years of endocrine therapy and during or within the first year of completing adjuvant ET [27].

Hormone receptor positivity was defined as ER or PR positivity with a percentage of immunohistochemistry (IHC)-positive nuclei of at least 1%. HER2 positivity was defined as a 3+ score at IHC or 2+ with amplified FISH, according to the 2018 ASCO-CAP guidelines [28]. HER2 zero was defined as HER score 0 and HER2 low as HER2 1+ or HER2 2+ with non-amplified FISH [29]. The Ki67 value was calculated using the percentage of nuclei that were MIB-1 positive [30].

Endpoints. The primary endpoint was PFS prediction, defined as the time from treatment initiation to disease progression, death from any cause or last follow-up determined by the attending physician based on imaging or clinical evaluation. Patients alive without disease recurrence were censored at the last follow-up visit. We further evaluated the capability of the model to predict early progression (EP) within 6 months. Outcome assessors were not blinded.

2.3. Statistical analysis

Clinicopathological characteristics at baseline were described using standard descriptive statistics. Continuous variables were reported as median and interquartile range (IQR) and evaluated Mann-Whitney test. Categorical variables were reported as frequency and percentage and assessed using a Chi-Square test or Fisher's exact test, as appropriate. Statistical analyses were performed using R Studio (ver 4.2.2) and Python (ver 3.10). All hypotheses were two-sided and $p < 0.05$ was considered statistically significant.

Sample size considerations. For PFS, we considered an R^2 of 0.4 and up to 15 candidate predictors and a median PFS of 30 months (corresponding to an event rate of 0.023). Considering 24 months as mean follow-up and 24 months as the timepoint, the minimum sample size required was 275 patients corresponding to 10.2 events per predictors and 153 outcome events in the training set.

Survival analysis. Survival estimates were evaluated by the Kaplan-Maier method. Differences between groups were assessed by the log-rank test. Median follow-up time was calculated by the reverse Kaplan-Meier method. Median survival and survival probabilities at specified time-points were reported, as well as absolute survival differences. Cox proportional hazard regression models were applied to estimate hazard ratios (HRs). The Efron method was used to handle ties in the Cox likelihood. The R packages survival and survminer were used for survival curves and diagnostic.

2.4. Model development and assessment

Pre-processing. The cohort was split in a train/internal validation set using an 70/30 split, stratified by institution. Continuous variables were scaled using z-score standardization and categorical were one-hot encoded. Missing data was imputed using the median value for continuous variable and the most frequent value for categorical variables.

Model development. Two ML models were trained using the scikit-survival package: regularized CoxNet regression (Cox proportional hazards model by incorporating elastic net regularization), and gradient boosting machines (GBM) (ensemble learning algorithm that builds predictive models by sequentially combining many weak learners decision trees). Hyperparameters tuning was performed using grid search and each model was evaluated using 5-fold cross-validation. Harrel C was used as the scoring metric. Both models were evaluated on the validation set. Model performance was evaluated with Harrell's C-index, Uno C-index, Brier score at specified time points and integrated Brier score, cumulative/dynamic AUC at specific time-points and mean AUC. 95% confidence intervals (CI) were computed via 1000 bootstrap replicates. Variable importance was assessed using model coefficients for CoxNet. Feature importance and SHapley Additive exPlanations (SHAP) values (with Kernel Estimator) were analysed for GBM. Permutation importance (30 permutations) was evaluated using Harrell C difference.

Risk stratification for survival. Risk stratification was performed on the train set using Gaussian Mixture Modelling (GMM) to define high- and low-risk groups based on predicted risk scores. A sensitivity analysis was conducted using the maximally selected rank statistics (MSRS). Cox regression estimated the corresponding hazard ratios (HRs).

Risk stratification for early progression. The model's ability to predict early progression at 6 months was assessed using the area under the receiver operating characteristic curve (AUROC) and precision-recall AUC (PR-AUC). The cut-off for 6 months binary prediction was calculated on the train set to maximize the F1 score. Decision curve analysis was performed.

3. Results

3.1. Cohort characteristics

From April 2017 to March 2025, 459 patients were included in the analysis, with a median follow-up time of 43.7 months (95% CI 39.6–48.3). We observed 253 events, with a median PFS of 29.3 months (95% CI 24.0–33.7) (Fig. S1).

3.2. Patients' characteristics

Median age at metastatic diagnosis was 61 years (IQR 52–71); the premenopausal was 25.1% and postmenopausal 74.3%; patients received AIs in 69.9% of cases and fulvestrant in 30.1% of cases. The HER2 receptor status was in 35.9 % HER2 0 and in 64.1% HER2-low. Most of the patients had a ductal histotype (74.7% ductal while 23.1% lobular). 69.9% of metastatic disease were considered hormone sensitive, 8.1% primary hormone resistant and 22% secondary hormone resistant. Median ER and PR expression were respectively 90 (IQR 85–90) and 40 (IQR 2–80). Median Ki67 value was 22.5% (IQR 15–35). 45.5% of patients received a primary diagnosis at metastatic stage while 53.4% patients received a previously diagnosis of early breast cancer. Prevalence of metastatic sites was distributed as follows: bone 68.2%, liver 27.9%, peritoneum 9.2%, brain 5.2%; while 18.3% of patient had bone only disease (Table 1).

After splitting, 321 patients were included in the training set, with 175 PFS events and a median PFS of 30.8 months (95%CI 25.2–35.7), and 138 in the validation set, with 78 PFS events and a median PFS of 23.0 months (95%CI 17.2–34.8) (Fig. S2). There was no significant difference between the two sets (p-value log-rank test = 0.17). The train set was used for exploratory feature selection and to train the ML models. The validation set was set aside and used only to test model predictions.

3.3. Model development and performance

In CoxNet regression, the best CV alfa was 0.017 (Fig. S3, Table S1), corresponding to a mean CV Harrell's C-index of 0.73 in the train set, and a Harrell's C of 0.74 (95% CI 0.67–0.80) in the validation set. The integrated Brier score was 0.16 and the mean AUC 0.78 (95%CI 0.69–0.85) (Fig. 1a). The GBM model (final hyperparameters after CV: learning rate = 1, maximum depth = 1, number of estimators = 16; Table S2) achieved a mean CV Harrell's C-index of 0.74 in the train set and a Harrell's C of 0.74 (95% CI 0.67–0.80) in the validation set. The integrated Brier score was 0.16, and the mean AUC 0.77 (95%CI 0.69–0.85) (Fig. 1b). Both models showed consistent estimates across time-points, and Uno's concordance index was comparable to Harrell's C-index, suggesting minimal impact of censoring on concordance assessment (Table S3).

3.4. Feature importance

The most important variables for CoxNet, were the presence of liver metastasis (0.43), Ki67 (0.24), primary hormone resistance (0.22) and the presence of metastasis in bone (0.19), peritoneum (0.19) and brain (0.18). The presence of bone-only disease was associated with better prognosis (−0.14) (Fig. 2a, Fig. S4). For GBM, feature importance analysis consistently identified metastatic sites - particularly liver (0.33), peritoneal (0.15), and brain (0.06) involvement - as the strongest predictors, alongside Ki67 (0.15), primary hormone resistance (0.09), PgR (0.08), and age at diagnosis (0.05) (Fig. 2b). Consistent results were observed for permutation importance (Fig. 2c–Tables S4–5) and SHAP variable contribution (Fig. 2d). Examples of single patients SHAP explanation are reported in Fig. S5.

Table 1
Patients' characteristics.

	Overall (n = 459)	Training (n = 321)	Validation (n = 138)	P-Value
Age (years)	61.8 [52.0–71.0]	61.8 [52.7–70.6]	61.9 [50.5–71.5]	0.893
Missing	0 (0.0)	0 (0.0)	0 (0.0)	
Menopausal status				
Premenopausal	115 (25.1)	74 (23.1)	41 (29.7)	0.314
Postmenopausal	341 (74.3)	245 (76.3)	96 (69.6)	
Missing	3 (0.7)	2 (0.6)	1 (0.7)	
ER	90.0 [85.0–90.0]	90.0 [85.0–90.0]	90.0 [85.0–90.0]	0.301
Missing	16 (3.5)	9 (2.8)	7 (5.1)	
PgR	40.0 [2.0–80.0]	40.0 [2.0–75.0]	40.0 [3.0–80.0]	0.586
Missing	19 (4.1)	11 (3.4)	8 (5.8)	
HER2 Status				0.655
Zero	165 (35.9)	118 (36.8)	47 (34.1)	
Low	294 (64.1)	203 (63.2)	91 (65.9)	
Ki67	22.5 [15.0–35.0]	22.0 [13.0–35.0]	25.0 [15.0–35.0]	0.513
Missing	29 (6.3)	20 (6.2)	9 (2.8)	
Histotype				0.498
Ductal	343 (74.7)	235 (73.2)	108 (78.3)	
Lobular	106 (23.1)	79 (24.6)	27 (19.6)	
Missing	10 (2.2)	7 (2.2)	3 (2.2)	
Brain metastases				0.645
No	433 (94.3)	302 (94.1)	131 (94.9)	
Yes	24 (5.2)	17 (5.3)	7 (5.1)	
Missing	2 (0.4)	2 (0.6)	0 (0.0)	
Liver metastases				0.161
No	329 (71.7)	237 (73.8)	92 (66.7)	
Yes	128 (27.9)	82 (25.5)	46 (33.3)	
Missing	2 (0.4)	2 (0.6)	0 (0.0)	
Bone metastases				0.166
No	144 (31.4)	93 (29.0)	51 (37.0)	
Yes	313 (68.2)	226 (70.4)	87 (63.0)	
Missing	2 (0.4)	2 (0.6)	0 (0.0)	
Peritoneal metastases				0.415
No	415 (90.4)	287 (89.4)	128 (92.8)	
Yes	42 (9.2)	32 (10.0)	10 (7.2)	
Missing	2 (0.4)	2 (0.6)	0 (0.0)	
Bone only disease				0.335
No	373 (81.3)	256 (79.8)	117 (84.8)	
Yes	84 (18.3)	63 (19.6)	21 (15.2)	
Missing	2 (0.4)	2 (0.6)	0 (0.0)	
Metastatic at diagnosis				0.591
No	245 (53.4)	176 (54.8)	69 (50.0)	
Yes	209 (45.5)	142 (44.2)	67 (48.6)	
Missing	5 (1.1)	3 (0.9)	2 (1.4)	
Hormone resistance				0.810
Hormone Sensitive	321 (69.9)	227 (70.7)	94 (68.1)	
Primary Resistant	37 (8.1)	26 (8.1)	11 (8.0)	
Secondary Resistant	101 (22.0)	68 (21.2)	33 (23.9)	

Mann-Whitney test was used for continuous variables. Chi square test was used for categorical variables.

3.5. Risk stratification

We calculated the GBM risk score for each patient in the train set (Fig. 3a), and using a GMM we identified 0.14 on the predicted risk score as the cut-off for risk stratification (Fig. 3b). Using this value, we observed a statistically significant difference between low and high-risk patients, both in the training and validation set. In the train set the HR was 4.14 (95%CI: 3.04–5.64, $p = 1.7 \times 10^{-19}$), and the median PFS was 58.1 months (95%CI 40-8-NE) in low risk and 13.7 months (95%CI 9.8–17.5) in high-risk patients (Fig. 3c). Consistent results were observed in the validation set, with an HR of 2.58, 95% CI: 1.65–4.03, $p = 3.3 \times 10^{-5}$). Median PFS was 34.8 (95%CI 24.0–52.4) in the low-risk

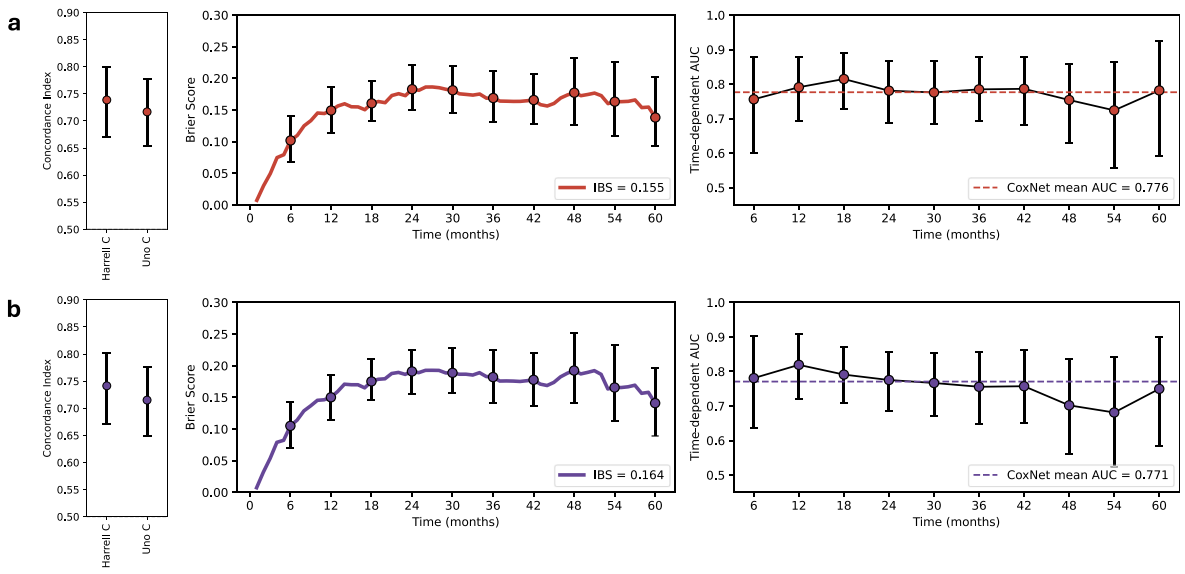

Fig. 1. Model performance comparison. a. CoxNet model (alpha = 0.0168). b. GBM (learning rate = 1.0, max depth = 1, n estimators = 16). From left to right: Harrell C Index and Uno C Index, Brier score at specific time points and integrated Brier score (time 0–60 months), and cumulative/dynamic AUC at specific time points and mean AUC (time 0–60 months). Error bars represent 95% CI.

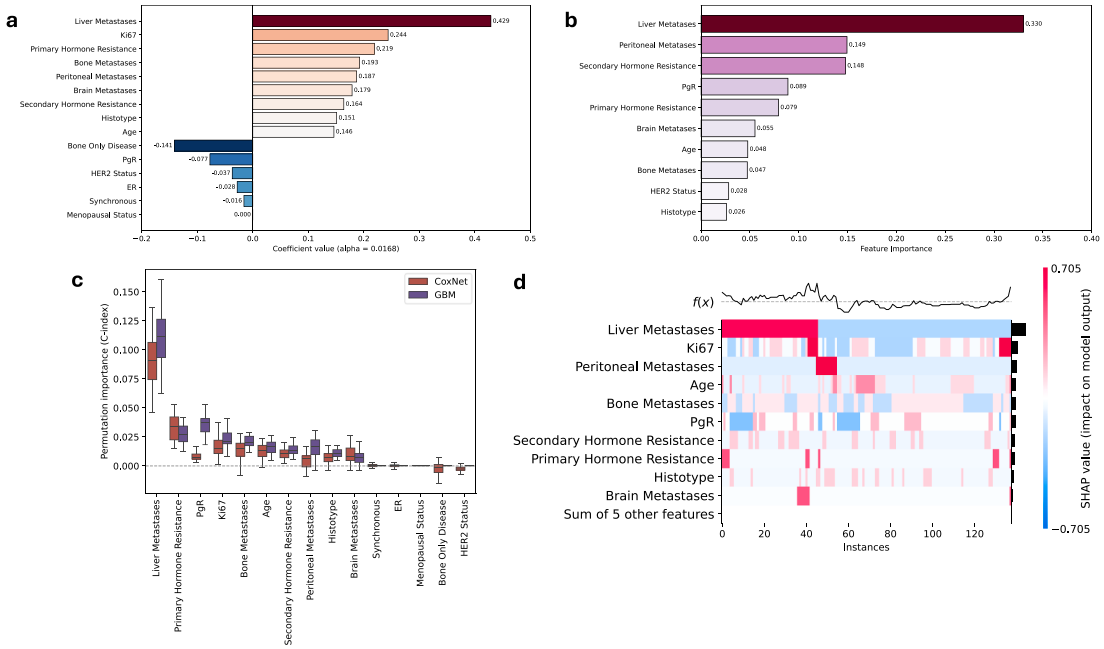

Fig. 2. Model explainability. a. CoxNet coefficients in the train set. b. GBM features importance in the train set. c. Comparison of permutated feature importance between the CoxNet and GBM model on validation set, based on the difference in Harrell C Index. Boxplot represents median and IQR range. d. Heatmap of SHAP feature contributions and risk scores in the validation set.

and 10.6 (95%CI 7.7–14.6) in the high-risk group (Fig. 3d). Consistent results were observed using MSRS to select the cut-off (Fig. S6).

3.6. Early progression risk

Overall, 18.7% patients progressed within six months. 124 had at least 6 months of follow-up in the validation set, and there were 19 events (15.3%). In the validation-set the model achieved an AUROC of 0.77 (95%CI 0.61–0.89) and a PR-AUC of 0.47 (0.28–0.70) (Fig. 4a–b). A cut-off of 0.20 for 6 the probability of 6 months progression achieved the best F1 score of 0.44 (Figs. S7–8), with fair calibration (Fig. S9), and the decision curve analysis showed model net benefit for this threshold on validation data (Fig. 4c).

4. Discussion

Currently, in daily practice for HR+/HER2− metastatic breast cancer (MBC), robust predictive factors to guide the choice of the best treatment for an individual patient are lacking. The exploratory analysis of the SONIA trial recently published confirmed that predictive tools for first-line endocrine-based regimens remain under investigation, with circulating tumor DNA (ctDNA) representing a promising but still

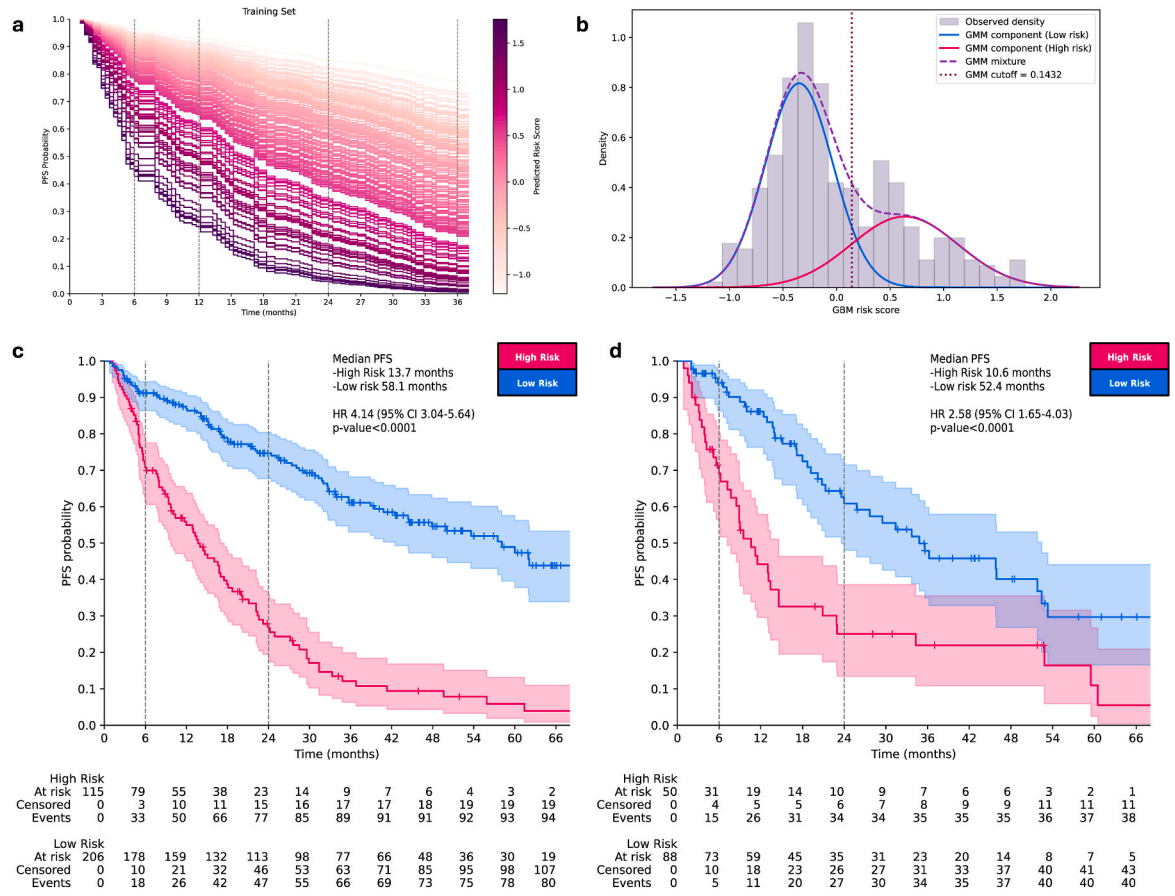

Fig. 3. Model predictions and risk stratification. a. Survival function for GBM, according to risk score. b. Distribution of risk score, with density of the 2 GMM components corresponding to risk groups. c. GBM Kaplan Maier survival curve in the train set, based on GMM cut-off. d. GBM Kaplan Maier survival curve in the validation set, based on GMM cut-off.

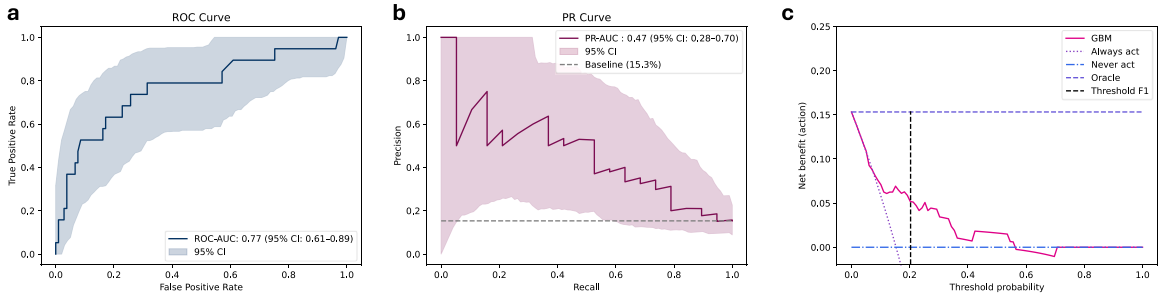

Fig. 4. Model performance and net benefit for early progression (<6 months). a. ROC curve and area under the ROC curve, with 95% CI in validation set b. Precision-recall curve and area under the PR curve, with 95% CI in validation set. c. Decision curve analysis and net benefit in the validation set. The cut-off maximizing F1 score in train set is reported.

unvalidated approach [31]. Recently, the INAVO120 trial, in patients with endocrine-resistant, PIK3CA-mutated, HR+/HER2− MBC, established the triplet regimen as a new first-line standard in this specific subgroup [32].

The present study supports the clinical applicability of GBM and CoxNet as robust tools for individualized risk stratification in HR + MBC patients treated with CDK4/6 inhibitors, based solely on baseline clinical parameters. We observed a statistically significant difference between low- and high-risk predicted groups, with an HR of 2.58 (95% CI: 1.65–4.03, $p = 3.3 \times 10^{-5}$) in the validation set. These models highlight the crucial prognostic role of metastatic site distribution and Ki-67 expression in the risk of progression.

In the Memorial Sloan Kettering translational database, a GBM algorithm was used to generate three models to predict CDK4/6 inhibitor PFS based on clinicopathological features (CF), genomic features (GF), and combined CF and GF (CGF). The HRs between good- and poor-risk groups were 1.95 (95% CI 1.5–2.5; p = 2.6e-7), 2.4 (95% C: 1.9–3.1; p = 2.0e-11), and 4.2 (95% CI 3.0–5.9; p = 2.8e-16) for the CF, GF, and CGF models, respectively. The most important features were tumor mutational burden (TMB), fraction of genome altered, TP53 alteration, fraction of genome with loss of heterozygosity (LOH), presence of liver metastases, adjuvant treatment-free interval <1 year, primary tumor grade 3, presence of visceral metastases, PR negativity, and whole-genome doubling [33]. In a second work of the same group evaluating

only clinicopathological features, the GBM model stratified patients into three risk groups with significantly different PFS between good-response and poor response (HR 2.55, 95% CI 2.0–3.2 p = 2.54e-14), highlighting disease-free interval, liver involvement, age at metastatic disease, adjuvant treatment-free interval, and progesterone receptor status as key predictors [34]. The results of our model are highly consistent with those reported in the previous data, identifying a comparable importance of variables. Considering only clinicopathological features, our model achieves the same stratification capacity, with a robust Harrell's C-index of 0.74 (95% CI: 0.67–0.80) in the validation set. We did not observe marked performance difference between GBM and CoxNet. Coherently with previous observations, this could be related to limited number of predictors and the moderate size of the dataset [35]. The GBM offers flexibility to capture non-linear variables and interactions, without prior specification.

Moreover our findings are consistent with a retrospective analysis of seven phase III trials of CDK4/6 inhibitors plus ET, where a random survival forest model predicted PFS with a C-index of 0.62, histologic grade being the strongest predictor followed by menopausal status [36]. Real-world data confirmed that PR-negative status, low ER expression (<30%), and histology grade 3 correlated with shorter PFS, with a combined subgroup (G3/PR−) showing a PFS of only 6.2 months (HR 3.86), suggesting that PFS estimates could be deduced from a relatively small number of existing pathological measures [5]. The presence of liver metastases remains a major adverse factor, with a median OS of approximately two years despite CDK4/6 inhibitors plus ET [37]. The analysis of the MD Anderson database showed no PFS or OS differences between ductal and lobular histotype [38], though peritoneal metastases (often associated with lobular disease) were predictive in our cohort (lobular 23.1%, peritoneal 9.2%). Brain metastases (only 5.2% of our patients' sample) are reported to be related to a biologically aggressive disease. In a non-randomized prospective phase II study, 58 patients with brain metastases were treated with abemaciclib in the median fourth line of therapy, with a poor intracranial response rate (ORR 5.2%) and a median OS of 12.5 months [39]. The prognostic role of HER2-low status remains debated, with conflicting results in the ET setting. Similarly, within our model, HER2-low did not demonstrate predictive significance.

Our data show that the model achieved an AUROC of 0.77 (95% CI: 0.61–0.89), in predicting early progression. In this regard, an observational longitudinal real-world study of first-line ET±CDK4/6 inhibitors identified Elastic Net as the best predictive model (AUC 0.73) for early progression (<6 months) [40]. While Using LASSO feature selection, delta-radiomic CT features further predicted early response with an AUC of 0.72 [41].

Recent evidence further supports the prognostic and predictive role of liquid biopsy approaches in HR+/HER2-metastatic breast cancer, with: 1) circulating tumor cells (CTCs) emerging as significant predictors of treatment outcome in MONARCH 2 [42,43], 2) ctDNA redefining the potential of non-invasive monitoring and molecular stratification [44], and 3) integrated CTC/ctDNA machine learning models providing proof-of-principle for endocrine resistance profiling [45]. While CTCs and ctDNA independently provide valuable insights into treatment outcomes and tumor dynamics, combining liquid biopsy data with machine learning increases the ability to profile endocrine resistance and thus enables more accurate risk stratification. The results of the SERENA-6 [46] and INAVO120 [32] trials highlight the increasing role of biomarker-guided strategies in HR+/HER2− MBC and provide an opportunity to put in the context the predictive capabilities of artificial intelligence models. For high-risk patients, intensified monitoring of ctDNA for ESR1 or early deployment of PI3K inhibition in the presence of PIK3CA mutations represent rational interventions to delay progression and overcome early endocrine resistance. Conversely, low-risk patients, may continue to benefit from standard CDK4/6-based strategies without the need for early intensification. This hypothesis could also influence the cost-effectiveness of molecular analyses.

Limitations should be acknowledged: being a retrospective analysis from two institutions the possibility of selection bias, measurement bias, missing data, or simplification of pathologic reports cannot be excluded. In our database, histological grade at the time of metastatic diagnosis was not available for most patients and therefore not considered; however, in other retrospective studies, this feature was identified as predictive in survival models using ML [33,36]. The lack of external validation may limit the generalizability of our findings and external confirmation is warranted before clinical implementation. Genomic alterations and ctDNA assessment were not included in the model, which may further improve predictive performance. This is coherent with our primary objective of develop a clinically applicable risk stratification model based on variables routinely available in clinical practice. In this light, to allow independent validation we provided a GitHub repository with step-by-step instructions to assess performance on external datasets, and to facilitate potential clinical implementation we developed a graphical interface to allow real-time assessment of patient risk, available at https://cdk46i-metastatic-bc-ml-jkc3ij9ut3wu6cc3nphjps.streamlit.app/. In this context, model predictions could be used to identify high-risk patients, which could be candidates for integration of molecular profiling (e.g., ctDNA or targeted sequencing), and enrolment in clinical trials evaluating intensified or biomarker-driven strategies. Conversely, in low-risk patients, therapy adjustments to balance efficacy with cumulative toxicities could be considered [47]. Escalation and de-escalation strategies remain hypothesis-generating and should be prospectively tested in risk-based clinical trials.

We aimed to analyse only objective features collected at baseline. All patients were included regardless of performance status, thus representing an unselected, real-world population. Combining results of survival models and adverse event models can provide a more complete picture for both patients and physicians.

The choice of defining early progression as occurring within six months is consistent with clinical practice. Although the literature on predicting EP to first-line treatment in HR+/HER2− MBC is limited, some retrospective studies using machine learning models have adopted a similar definition [40,48].

Retrospective real-world studies have shown that germline BRCA1/2 pathogenic variants are independently associated with poor outcomes in patients with advanced breast cancer treated with CDK4/6 inhibitors plus ET [49,50]. In our cohort, 4.7% of all patients presented with a documented germline BRCA1/2 pathogenic variant. Only 125 of 459 patients (27%) underwent genetic testing; among those tested, 22 of 125 (17.6%) carried a germline pathogenic variant. Additionally, 3.2% of all patients have ESR1 or PIK3CA/AKT/PTEN pathway mutations, but only 130 patients were tested. Due to the small size and heterogeneity of the tested subgroup, these variables were not analysed.

Finally, 48.8%, 41.8%, and 9.3% of patients received palbociclib, ribociclib, and abemaciclib, respectively, in our dataset. This variable was not considered in the predictive evaluation, as treatment choice largely depended on the timing of drug availability in clinical practice and physician decision-making.

Overall, integration of AI-driven risk stratification with biomarker-directed therapeutic interventions holds promise to define a future paradigm in HR+/HER2− MBC management: a precision oncology approach where clinical, genomic, and real-time liquid biopsy data are combined to guide individualized therapy sequencing and a personalized escalation or de-escalation of therapy based on biologically informed prediction. Further studies are required to evaluate the impact of such multimodal approaches on survival outcomes.

CRediT authorship contribution statement

Sergio Pannunzio: Writing – review & editing, Writing – original draft, Supervision, Investigation, Conceptualization. **Luca Mastrantoni:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Noemi**

Maliziola: Supervision, Methodology, Investigation, Data curation, Conceptualization. **Letizia Pontolillo:** Supervision, Methodology, Investigation, Data curation, Conceptualization. **Giovanna Garufi:** Supervision, Methodology, Investigation, Data curation, Conceptualization. **Elena Di Monte:** Investigation, Data curation. **Alessandra Emiliani:** Investigation, Data curation. **Margherita Sgambato:** Investigation, Data curation. **Anna Cardillo:** Investigation, Data curation. **Antonella Palazzo:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Armando Orlandi:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Formal analysis. **Giampaolo Tortora:** Validation, Supervision, Conceptualization. **Emilio Bria:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Preliminary results of this analysis were presented as a poster at the ESMO annual congress 2025

Poster 516P 'Gradient Boosting Machine (GBM) for Predicting Progression-Free Survival (PFS) in Hormone-Receptor Positive (HR+) Metastatic Breast Cancer (mBC) receiving upfront CDK4/6 Inhibitors (CDK4/6i).'

Generative AI statement

During the preparation of this work, the authors used ChatGPT (OpenAI) to improve language clarity. The authors reviewed and edited all content and take full responsibility for the publication.

Funding

None declared.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: AO has declared consulting fees/advisory role for Novartis, Roche, Eli-Lilly, Amgen, Daiichi Sankyo, travel and accommodation by Daiichi Sankyo, Novartis, Roche, Pfizer. AP has declared consulting fees/advisory role for Amgen, MSD, Novartis, travel and accommodation by Pfizer. GT is supported by funds of Ministero della Salute (Ricerca Corrente 2025). EB is supported by Institutional funds of Università Cattolica del Sacro Cuore (UCSC-projects D1), by the AIRC under Investigator Grant (IG) No. IG20583 and the Italian Ministry of Health Ricerca Corrente 2025. EB received speakers' and travels' fee from MSD, Astra-Zeneca, Celgene, Pfizer, Helsinn, Eli-Lilly, BMS, Novartis, and Roche. EB received institutional research grants from Astra-Zeneca, Roche. All other authors declare no financial or non-financial competing interests.

Acknowledgements

Ministry of Health-Ricerca Corrente 2026 MoH-Rc2026.

Appendix A. Supplementary data

Supplementary data to this article can be found athttps://doi.org/10.1016/j.breast.2026.104715.

Data availability

The code to perform external independent validation and to reproduce the validation results of this work is available at https://github.com/LucaMastrantoni/cdk46i-metastatic-bc-ml. The graphical interface for model predictions on single patients is available at https://cdk46i-metastatic-bc-ml-jkc3ij9ut3wu6cc3nphjps.streamlit.app/.

References

[1] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021. https://doi.org/10.3322/caac.21660.
[2] Cardoso F, Paluch-Shimon S, Senkus E, et al. 5th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 5). Ann Oncol 2020;31(12):1623–49. https://doi.org/10.1016/j.annonc.2020.09.010.
[3] McAndrew NP, Finn RS. Clinical review on the management of hormone receptor–positive metastatic breast cancer. JCO Oncol Pract 2022;18(5):319–27. https://doi.org/10.1200/op.21.00384.
[4] Gao JJ, Cheng J, Bloomquist E, et al. CDK4/6 inhibitor treatment for patients with hormone receptor-positive, HER2-negative, advanced or metastatic breast cancer: a US Food and Drug Administration pooled analysis. Lancet Oncol 2020;21(2):250–60. https://doi.org/10.1016/S1470-2045(19)30804-6.
[5] Witkiewicz AK, Schultz E, Wang J, et al. Determinants of response to CDK4/6 inhibitors in the real-world setting. npj Precis Oncol 2023;7(1):1–17. https://doi.org/10.1038/s41698-023-00438-0.
[6] Condorelli R, Spring L, O'Shaughnessy J, et al. Polyclonal RB1 mutations and acquired resistance to CDK 4/6 inhibitors in patients with metastatic breast cancer. Ann Oncol 2018;29(3):640–5. https://doi.org/10.1093/annonc/mdx784.
[7] André F, Su F, Solovieff N, et al. Pooled ctDNA analysis of MONALEESA phase III advanced breast cancer trials. Ann Oncol 2023:1–12. https://doi.org/10.1016/j.annonc.2023.08.011.
[8] Finn RS, Liu Y, Zhu Z, et al. Biomarker analyses of response to cyclin-dependent kinase 4/6 inhibition and endocrine therapy in women with treatment-naïve metastatic breast cancer. Clin Cancer Res 2020;26(1):110–21. https://doi.org/10.1158/1078-0432.CCR-19-0751.
[9] Turner NC, Liu Y, Zhu Z, et al. Cyclin E1 expression and palbociclib efficacy in previously treated hormone receptor-positive metastatic breast cancer. J Clin Oncol 2019;37(14):1169–78. https://doi.org/10.1200/JCO.18.00925.
[10] Gong X, Litchfield LM, Webster Y, et al. Genomic aberrations that activate D-type cyclins are associated with enhanced sensitivity to the CDK4 and CDK6 inhibitor abemaciclib. Cancer Cell 2017;32(6):761–776.e6. https://doi.org/10.1016/j.ccell.2017.11.006.
[11] Álvarez-Fernández M, Malumbres M. Mechanisms of sensitivity and resistance to CDK4/6 inhibition. Cancer Cell 2020;37(4):514–29. https://doi.org/10.1016/j.ccell.2020.03.010.
[12] Li Z, Razavi P, Li Q, et al. Loss of the FAT1 tumor suppressor promotes resistance to CDK4/6 inhibitors via the hippo pathway. Cancer Cell 2018;34(6):893–905.e8. https://doi.org/10.1016/j.ccell.2018.11.006.
[13] Formisano L, Lu Y, Servetto A, et al. Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. Nat Commun 2019;10(1):1–14. https://doi.org/10.1038/s41467-019-09068-2.
[14] Asghar US, Kanani R, Roylance R, Mittnacht S. Systematic review of molecular biomarkers predictive of resistance to CDK4/6 inhibition in metastatic breast cancer. JCO Precis Oncol 2022;(6):1–13. https://doi.org/10.1200/po.21.00002.
[15] Stanciu IM, Parosanu AI, Orlov-Slavu C, et al. Mechanisms of resistance to CDK4/6 inhibitors and predictive biomarkers of response in HR+/HER2-Metastatic breast cancer—A review of the literature. Diagnostics 2023;13(5). https://doi.org/10.3390/diagnostics13050987.
[16] Fribbens C, O'Leary B, Kilburn L, et al. Plasma ESR1 mutations and the treatment of estrogen receptor-positive advanced breast cancer. J Clin Oncol 2016;34(25):2961–8. https://doi.org/10.1200/JCO.2016.67.3061.
[17] Prat A, Chaudhury A, Solovieff N, et al. Correlative biomarker analysis of intrinsic subtypes and efficacy across the MONALEESA phase III studies. J Clin Oncol 2021;39(13):148–1467. https://doi.org/10.1200/JCO.20.02977.
[18] Taghizadeh E, Heydarheydari S, Saberi A, JafarpoorNesheli S, Rezaeijo SM. Breast cancer prediction with transcriptome profiling using feature selection and machine learning methods. BMC Bioinf 2022;23(1):1–9. https://doi.org/10.1186/s12859-022-04965-8.
[19] Subbiah V. The next generation of evidence-based medicine. Nat Med 2023;29(1):49–58. https://doi.org/10.1038/s41591-022-02160-z.
[20] Noman SM, Fadel YM, Henedak MT, et al. Leveraging survival analysis and machine learning for accurate prediction of breast cancer recurrence and metastasis. Sci Rep 2025;15(1):1–16. https://doi.org/10.1038/s41598-025-87622-3.
[21] Lo Gullo R, Eskreis-Winkler S, Morris EA, Pinker K. Machine learning with multiparametric magnetic resonance imaging of the breast for early prediction of response to neoadjuvant chemotherapy. Breast 2020;49:115–22. https://doi.org/10.1016/j.breast.2019.11.009.
[22] Sammut SJ, Crispin-Ortuzar M, Chin SF, et al. Multi-omic machine learning predictor of breast cancer therapy response. Nature 2022;601(7894):623–9. https://doi.org/10.1038/s41586-021-04278-5.
[23] Mastrantoni L, Garufi G, Giordano G, et al. Accessible model predicts response in hormone receptor positive HER2 negative breast cancer receiving neoadjuvant chemotherapy. npj Breast Cancer 2025;11(1). https://doi.org/10.1038/s41523-025-00727-w.
[24] Pfob A, Sidey-Gibbons C, Rauch G, et al. Intelligent vacuum-assisted biopsy to identify breast cancer patients with pathologic complete response (ypT0 and ypN0)

after neoadjuvant systemic treatment for omission of breast and axillary surgery. J Clin Oncol 2022;40(17):1903–15. https://doi.org/10.1200/JCO.21.02439.
[25] Kuerer HM, Smith BD, Krishnamurthy S, et al. Eliminating breast surgery for invasive breast cancer in exceptional responders to neoadjuvant systemic therapy: a multicentre, single-arm, phase 2 trial. Lancet Oncol 2022;23(12):1517–24. https://doi.org/10.1016/S1470-2045(22)00613-1.
[26] Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. Br Med J 2024. https://doi.org/10.1136/bmj-2023-078378.
[27] Cardoso F, Senkus E, Costa A, et al. 4th ESO-ESMO international consensus guidelines for advanced breast cancer (ABC 4). Ann Oncol 2018;29(8):1634–57. https://doi.org/10.1093/annonc/mdy192.
[28] Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mcshane LM, Dowsett M. HER2 testing in breast cancer : American society of clinical oncology/college of American pathologists clinical practice guideline focused update summary, vol. 14; 2023. p. 437–42. https://doi.org/10.1200/JOP.18.00206.
[29] Modi S, Jacot W, Yamashita T, et al. Trastuzumab deruxtecan in previously treated HER2-Low advanced breast cancer. N Engl J Med 2022;387(1):9–20. https://doi.org/10.1056/nejmoa2203690.
[30] Dowsett M, Nielsen TO, A'Hern R, et al. Assessment of Ki67 in breast cancer: recommendations from the international Ki67 in breast cancer working group. J Natl Cancer Inst 2011;103(22):1656–64. https://doi.org/10.1093/jnci/djr393.
[31] Jongbloed EM, Wortelboer N, de Weerd V, et al. Early versus deferred use of CDK4/6 inhibitors in advanced breast cancer: circulating tumor DNA analysis of a randomized phase 3 trial. Nat Med 2025;2. https://doi.org/10.1038/s41591-025-03935-w.
[32] Jhaveri KL, Im S-A, Saura C, et al. Overall survival with inavolisib in PIK3CA -Mutated advanced breast cancer. N Engl J Med 2025;393(2):151–61. https://doi.org/10.1056/nejmoa2501796.
[33] Moiso E, Ferraro E, Cabel L, et al. Abstract GS3-09: multimodal integration of real world clinical and genomic data for the prediction of CDK4/6 inhibitors outcomes in patients with HR+/HER2- metastatic breast cancer. Clin Cancer Res 2025;31 (12_Supplement):GS309. https://doi.org/10.1158/1557-3265.sabcs24-gs3-09. GS3-09.
[34] Moiso E, Ferraro E, Cabel L, et al. Predicting CDK4/6 inhibitors outcomes in pts with HR+/HER2- metastatic breast cancer: a machine learning approach. J Clin Oncol 2024;42(16_suppl):1074. https://doi.org/10.1200/jco.2024.42.16_suppl.1074. 1074.
[35] Mastrantoni L, Garufi G, Giordano G, et al. Comparison of machine learning and deep learning models for survival prediction in early-stage hormone receptor-positive/HER2-negative breast cancer receiving neoadjuvant chemotherapy. ESMO Real World Data Digit Oncol 2025;10(C):100184. https://doi.org/10.1016/j.esmorw.2025.100184.
[36] Mason J, Gong Y, Amiri-Kordestani L, et al. Model development of CDK4/6 predicted efficacy in patients with hormone receptor–positive, human epidermal growth factor receptor 2–Negative advanced or metastatic breast cancer. JCO Clin Cancer Informatics 2021;(5):758–67. https://doi.org/10.1200/cci.21.00025.
[37] Ji L, Cheng L, Zhu X, Gao Y, Fan L, Wang Z. Risk and prognostic factors of breast cancer with liver metastases. BMC Cancer 2021;21(1):1–15. https://doi.org/10.1186/s12885-021-07968-5.
[38] Mouabbi JA, Raghavendra AS, Bassett RL, Hassan A, Tripathy D, Layman RM. Histology-based survival outcomes in hormone receptor-positive metastatic breast cancer treated with targeted therapies. npj Breast Cancer 2022;8(1):1–7. https://doi.org/10.1038/s41523-022-00499-7.
[39] Tolaney SM, Sahebjam S, Le Rhun E, et al. A phase II study of abemaciclib in patients with brain metastases secondary to hormone receptor-positive breast cancer. Clin Cancer Res 2020;26(20):5310–9. https://doi.org/10.1158/1078-0432.CCR-20-1764.
[40] Ribelles N, Jerez JM, Rodriguez-Brazzarola P, et al. Machine learning and natural language processing (NLP) approach to predict early progression to first-line treatment in real-world hormone receptor-positive (HR+)/HER2-negative advanced breast cancer patients. Eur J Cancer 2021;144:224–31. https://doi.org/10.1016/j.ejca.2020.11.030.
[41] Khorrami M, Viswanathan VS, Reddy P, et al. Radiomic predicts early response to CDK4/6 inhibitors in hormone receptor positive metastatic breast cancer. npj Breast Cancer 2023;9(1):1–9. https://doi.org/10.1038/s41523-023-00574-7.
[42] Gerratana L, Kocherginsky M, Davis AA, et al. Circulating tumor cells prediction in hormone receptor positive HER2-Negative advanced breast cancer: a retrospective analysis of the MONARCH 2 trial. Oncologist 2024;29(2):123–31. https://doi.org/10.1093/oncolo/oyad293.
[43] Cristofanilli M, Pierga JY, Reuben J, et al. The clinical use of circulating tumor cells (CTCs) enumeration for staging of metastatic breast cancer (MBC): international expert consensus paper. Crit Rev Oncol Hematol 2019;134(December 2018): 39–45. https://doi.org/10.1016/j.critrevonc.2018.12.004.
[44] Malik S, Zaheer S. The impact of liquid biopsy in breast cancer: redefining the landscape of non-invasive precision oncology. J Liq Biopsy 2025;8(April):100299. https://doi.org/10.1016/j.jlb.2025.100299.
[45] Gerratana L, Kocherginsky M, Davis AA, et al. Circulating tumor cells prediction in hormone receptor positive HER2-Negative advanced breast cancer : a retrospective analysis of the MONARCH 2 trial (July). 2023. p. 1–9.
[46] Bidard F-C, Mayer EL, Park YH, et al. First-line camizestrant for emerging ESR1 -Mutated advanced breast cancer. N Engl J Med 2025:569–80. https://doi.org/10.1056/nejmoa2502929.
[47] Mastrantoni L, Orlandi A, Palazzo A, et al. The likelihood of being helped or harmed as a patient-centred tool to assess cyclin dependent kinase 4/6 inhibitors clinical impact and safety in metastatic breast cancer: a systematic review and sensitivity-analysis. eClinicalMedicine 2023;56(January):1–13. https://doi.org/10.1016/j.eclinm.2023.101824.
[48] Lee JS, Yost SE, Li SM, et al. Genomic markers of CDK 4/6 inhibitor resistance in hormone receptor positive metastatic breast cancer. Cancers (Basel) 2022;14(13): 1–13. https://doi.org/10.3390/cancers14133159.
[49] Collins JM, Nordstrom BL, McLaurin KK, et al. A real-world evidence study of CDK4/6 inhibitor treatment patterns and outcomes in metastatic breast cancer by germline BRCA mutation status. Oncol Ther 2021;9(2):575–89. https://doi.org/10.1007/s40487-021-00162-4.
[50] Bottosso M, Zurlo C, Miglietta F, et al. Impact of germline BRCA1/2 status on outcomes for patients with HR+/HER2- metastatic breast cancer treated with CDK4/6 inhibitors: a systematic review and meta-analysis. Breast 2025;83(April): 104544. https://doi.org/10.1016/j.breast.2025.104544.