PERSONALIZED MEDICINE IN ADVANCED BREAST CANCER: AI-DRIVEN GENOMIC TEST FOR CDK4/6 TREATMENT RESPONSE PREDICTION.

1134P

AUTHORS AND AFFILIATIONS: Zawadzki Paweł[1,2], Woźna Alicja[1,2], Piernik Maciej[1,3], Gniot Michał[1,4], Pacewicz Klaudia[1], Sielski Dawid[1], Wojtaszewska Marzena[1,5], Brzeziński Dariusz[1,3], Rosiak Jacek[1], Stępień Rafał[6], Dąbrowski Maciej[1], Kraszewski Andrzej[1], Majer-Burman Weronika[1], Sztromwasser Paweł[1]

[1]MNM Bioscience, Inc, 16192 Coastal Highway, Lewes, DE 19958; [2]Faculty of Physics, Adam Mickiewicz University, Uniwersytetu Poznańskiego Street 2, 61-614 Poznan, Poland; [3]Institute of Computing Science, Poznan University of Technology, Piotrowo Street 2, 60-965 Poznan, Poland; [4]Department of Hematology and Bone Marrow Transplantation, Poznan University of Medical Sciences, Szamarzewskiego Street 84, 60-569, Poznan, Poland.; [5]Central Clinical Hospital of the Internal Affairs in Warsaw, Woloska Street 137, 02-507 Warsaw, Poland; [6]Biostatistics Group, Department of Genetics, Wrocław University of Environmental and Life Sciences, Norwida Street 25, 50-375 Wrocław, Poland

INTRODUCTION

Targeted therapy, using CDK4/6 inhibitors has been proven to be effective in patients with advanced breast cancer. The FDA has approved three CDK4/6 inhibitors (CDK4/6i; palbociclib, abemaciclib, and ribociclib) for patients with HR-positive (HR+)/HER2-negative (HER2-) breast cancer. Despite their apparent effectiveness, a significant proportion of patients do not benefit from the treatment [1, 2]. Here, we show that the amplification levels in 8 genomic regions demonstrate superior predictive power and enable accurate patient stratification.

The CDK4/6iDx multi-biomarker test
selects patients who will respond to CDK4/6i therapy.

METHODS

Recent studies have failed to deliver unambiguous genomic biomarker 'candidates' for CDK4/6-targeted therapy [2, 3]. Consisted with that, our efforts to identify genomic mutations as CDK4/6i predictive biomarkers were unsuccessful. In order to identify biomarkers for CDK4/6i, we have developed an agnostic AI-powered methodology able to identify genomic features with the highest predictive power. First, we have built BASE (Biomarker-Auto Search Engine), which can be trained to identify distribution-based genomic features, discriminating CDK4/6i responders from non-responders. Mining of 170 Whole Genome Sequencing (WGS) data from HR+/HER2- tumor biopsies resulted in a list o biomarker 'candidates'. Next, we created Random Similarity Forest (RSF) algorithm [4] which, in contrast to the typically used Random Forest (RF) approach, allows to build predictive models based not only on binary mutational status but also on the distributions of genomic changes. Together, BASE and RSF were used in building a **predictive model - CDK4/6iDx**, where the level of copy number alterations (CNAs) within 8 genomic regions has the highest discriminative power to select breast cancer patients who will respond to CDK4/6i treatment (Figure 1).

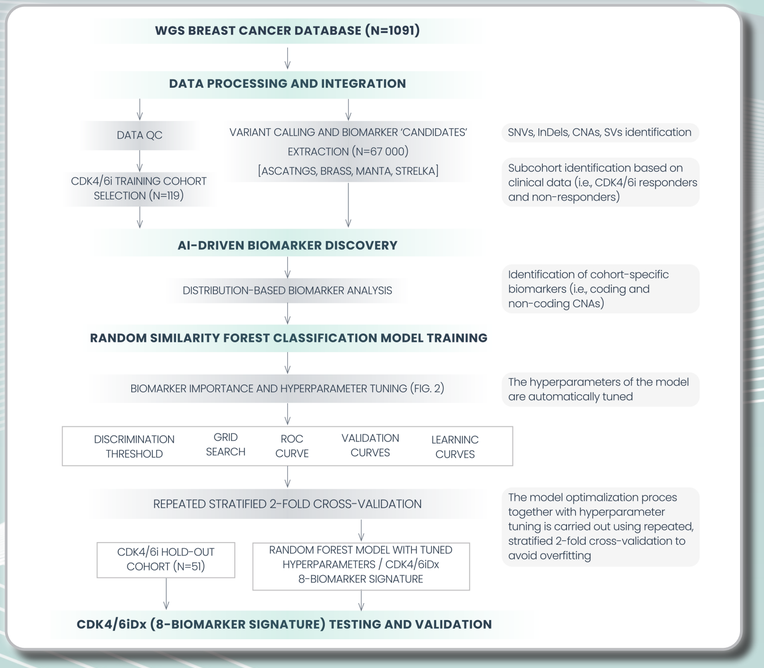


FIGURE 1. Analytical workflow. WGS data were processed and integrated with a proprietary pipeline, which provided variant calling, followed by BASE distribution-based biomarker extraction. RSF methodology was used for CDK4/6iDx classifier model training.

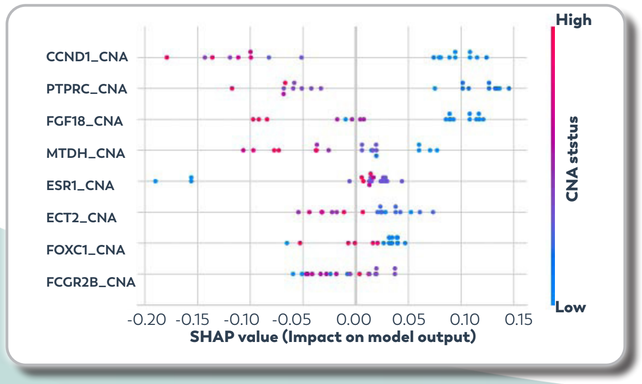


FIGURE 2. Ranking of eight genomic biomarkers using SHapley Additive exPlanation (SHAP). SHAP plot illustrates the association of the biomarkers' CNA status (red instances - high values, blue instances - low/normal values) with the impact on the Random Similarity Forest model output. Negative SHAP values are associated with resistance, positive values with sensitivity. Biomarkers are ranked in descending order of importance (impact on model output magnitude).

RESULTS

The AI-powered methodology, developed to discover novel biomarkers, identified 8 genomic regions – CNA hotspots containing numerous genes related to the cell cycle or cellular signal transduction, which predict the CDK4/6i response. We emphasize, that from a diagnostic point of view, the state of copy number alterations from a defined hotspot is uniform and can be analyzed interchangeably, therefore genes with known association to tumor biology were selected for the model training (Figure 2). **The predictive power of an interdependent biomarker collection (biomarker signature) is higher than that built on the basis of a single biomarker** (Figure 3).

The final test, conducted on a separate hold-out dataset of patient samples, confirmed the high predictive performance of CDK4/6iDx (Table 1).

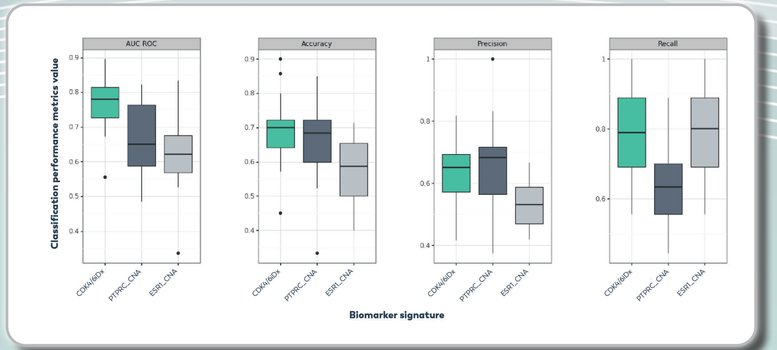
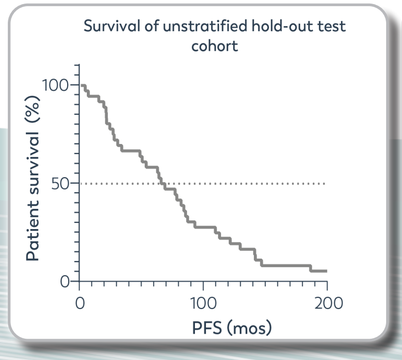


FIGURE 3. Computing cross-validation performance metrics for the CDK4/6iDx biomarker signature and two single-gene biomarkers. In the Random Forest model cross-validation, the CDK4/6iDx biomarker signature has more discriminative power than the single biomarkers, e.g., *PTPRC_CNA* and *ESR1_CNA* in the CDK4/6iDx signature, respectively. AUC ROC (Area Under the Receiver Operating Characteristics).

Retrospective validation on WGS hold-out test data has demonstrated that CDK4/6iDx has a great power of discriminating responders from non-responders (Figure 4). The PFS of CDK4/6i responders was longer, reaching 79 months vs. 58 months in non-responders. **CDK4/6iDx accurately stratifies advanced HR+/HER2- breast cancer patients,** allowing non-responders to evade serious adverse events. Preliminary results indicate that CDK4/6iDx biomarker signature is also present in other subtypes of breast cancer (HER2+, TNBC) as well as other cancer types (ovarian cancer).

TABLE 1. CDK4/6iDx performance on hold-out dataset.

	Accuracy	Precision	Recall	AUC ROC
RES	0.74	0.70	0.78	0.86



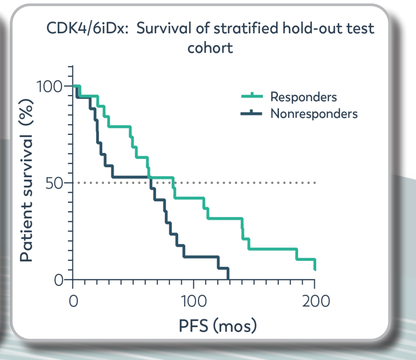


FIGURE 4. Kaplan-Meier curves for unstratified (left) and patients in-silico stratified by CDK4/6iDx (right), receiving CDK4/6-targeted therapy as second-line treatment. The PFS benefit was statistically significant (log-rank test, p=0.005).

CONCLUSIONS:

- The CDK4/6iDx is a multi-biomarker genomic stratification tool dedicated for patients with advanced HR+/HER2- breast cancer, with the potential to be extended to other indications.
- The CDK4/6iDx test consists of eight biologically important features – CNA hotspots, with recognized association to tumor biology.
- **The results demonstrate that a significant proportion of patients could avoid exposure to ineffective treatment with CDK4/6 inhibitors and the serious adverse events that come with it.**
- Validation of CDK4/6iDx is in progress. With more data available, the performance of the model will likely improve.

CONFLICT OF INTEREST

AW and PZ are share owners in the company MNM Bioscience Inc. 16192 Coastal Highway, Lewes, DE 19958 and are employed at Adam Mickiewicz University in Poznań. Other authors declare no conflict of interests.

REFERENCES:

[1] Im, S. et al. Overall Survival with Ribociclib plus Endocrine Therapy in Breast Cancer. N. Engl. J. Med. 381, 307-316 (2019).
[2] Rozenblit, M. et al. Patterns of treatment with everolimus exemestane in hormone receptor-positive HER2-negative metastatic breast cancer in the era of targeted therapy. Breast Cancer Res. 23, 1-10 (2021).
[3] Schoninger, S.F. & Blain, S.W. The Ongoing Search for Biomarkers of CDK4/6 Inhibitor Responsiveness in Breast Cancer. Mol. Cancer Ther. 19, 3-12 (2021).
[4] Piernik M., Brzezinski D., Random Similarity Forests, (in review)
[5] Jaksik, R., Rosiak, J., Zawadzki, P. & Sztromwasser, P. Accuracy of somatic variant detection workflows for whole genome sequencing experiments. bioRxiv (2021).