REVIEW

The role of artificial intelligence integrating multi-omics in breast cancer

Raquel Gómez-Bravo [a,b,1,*], Benjamín Walbaum [a,b,c,d,1], Elia Segui [a,b,e,f], Montserrat Muñoz [a,b,f]

[a] *Translational Genomics and Targeted Therapies in Solid Tumors, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Barcelona, Spain*
[b] *Cancer Institute and Blood Diseases, Hospital Clinic de Barcelona, Barcelona, Spain*
[c] *Department of Medical Oncology, Pontificia Universidad Catolica Santiago, Chile*
[d] *Faculty of Medicine, Pontificia Universidad Católica, Santiago, Chile*
[e] *Breast Oncology Center, Dana-Farber Cancer Institute, Boston, MA, USA*
[f] *SOLTI Breast Cancer Research Group, Barcelona, Spain*

Abstract   In an era of precision oncology, genomic testing plays a crucial role in the management of breast cancer. A variety of complex techniques for germline, somatic, and gene expression testing are routinely used as part of our clinical practice. However, challenges remain in both interpreting genomic data and in the ever-expanding breadth of available tumor information. Artificial intelligence (AI), specifically machine learning and deep learning models, can create and facilitate the interpretation of complex genetic data, predict patient outcomes, and personalize treatment plans. Herein, we present a review of the current role of AI integrating multi-omics in BC.
© 2025 SESPM. Published by Elsevier España, S.L.U. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

El papel de la Inteligencia Artificial en la integración de la multiómica en cáncer de mama

Resumen   En la era de la medicina de precisión, las pruebas genómicas desempeñan un papel crucial en el tratamiento del cáncer de mama (CM). En la práctica clínica se utilizan de manera habitual una variedad de técnicas complejas tanto para línea germinal como somática y de expresión genética. Sin embargo, persisten desafíos en la interpretación de los datos genómicos dada su complejidad creciente. La inteligencia artificial (IA) y, específicamente, los modelos de

---

* Corresponding author.
 E-mail addresses: gomezbravoraquel@gmail.com, (R. Gómez-Bravo), mmunoz@clinic.cat (M. Muñoz).
[1] Both authors contributed equally.

aprendizaje automático y aprendizaje profundo, tienen la capacidad de crear y facilitar la interpretación de datos genéticos complejos, predecir los resultados de los pacientes y personalizar los planes de tratamiento. Aquí, presentamos una revisión del papel actual de la IA en la integración de multiómicas en CM.

Introduction

The use of genomic testing has become a cornerstone in the management of breast cancer (BC), as it plays a critical role in patient care. These tests encompass a wide range of techniques and serve different purposes, including genomic screening for germline pathogenic variants that predispose to cancer development, to the identification of potential somatic mutations that act as "drivers" and hence are druggable targets[1]. Germline testing helps identify inherited cancer risks, while somatic testing focuses on acquired mutations. Both help improve prognostic assessments and enable the identification of therapeutic targets within an ever-expanding arsenal of therapies. However, breast cancer is a complex disease, often driven by multiple genomic features rather than a single-driver mutation[2]. While individual DNA alterations in tumor cells can be clinically useful, refining patients' prognosis and treatment outcomes may require additional biological information. Phenotypic characterization through multi-gene RNA-based expression analysis offers valuable insights by detecting the prognostic and predictive molecular intrinsic subtypes[3,4]. Still, with a growing array of "omic" technologies that provide a deeper understanding of cancer biology, an integrative multi-omics approach is necessary to generate a comprehensive analysis that better describes the complexity of biological systems behind cancer[5].

Around 5−10% of patients with BC carry hereditary pathogenic variants, hereafter defined as germline mutations, that drive the disease. The identification of these mutations in *BRCA or BRCA2* genes has implications beyond surgical decision-making, as they can inform treatment strategies involving target therapies such as poly(ADP-ribose) polymerase inhibitors (PARPi)[6]. Similarly, in patients with advanced hormone receptor-positive (HR+) BC, the detection of *PI3K/AKT* pathway mutations in tumor samples or through liquid biopsy for circulating tumor DNA (ctDNA) can guide the use of targeted therapies[7,8]. Patients with *ESR1* mutations may benefit from endocrine therapy with selective estrogen receptor degraders (SERDs), further optimizing treatment outcomes[9].

Gene expression platforms (GEPs) have been shown to help select patients with different prognoses who may benefit from chemotherapy or even require extended endocrine therapy in early-stage disease[10,11]. These techniques have been rapidly integrated into clinical practice. Additionally, pathology, including widely available basic tests such as immunohistochemistry (IHC) biomarkers, remains central in BC initial assessment[12,13]. However, with new drugs being approved, biomarkers are needed.

With the expanding availability of technologies that enable an in-depth understanding of the tumoral genome, epigenome, transcriptome, proteome, and metabolism, a single omics analysis provides only a partial view of tumoral biology. Thus, it is necessary to adopt an integrative, multi-omics approach, one that can even include unstructured data that clinicians would normally discard, to fully offer personalized cancer care.

In recent years, integrating artificial intelligence (AI) into genomics has transformed BC care. AI technologies, specifically machine learning (ML), a subfield of AI, have enhanced the ability to interpret complex genetic data, predict patient outcomes, and personalize treatment plans (Fig. 1). ML models combine previously defined features to make predictions[14]. As a result, ML models are usually simpler and can be used to predict the likelihood of recurrence based on a limited set of information like age, tumor size, and basic biomarkers previously validated as prognostic. Deep learning (DL) is an advanced subset of ML that mimics how the human brain works by learning from large amounts of data without needing manually defined features. DL uses neural networks with multiple layers to automatically identify patterns and make decisions based on both structured and unstructured data. This autonomous feature-learning capability makes DL a promising tool in cancer treatment and research[15].

Understanding the applications of AI in genomics is essential to managing the increasing volume and complexity of data. This review highlights the role of AI in integrating multi-omics data and addressing challenges in BC care.

Genomics in breast cancer

Genomics studies have steadily gained ground in BC, particularly germline testing, part of the initial evaluation for patients with certain clinical risk characteristics. Considering its predictive role for treatments with PARPi, germline testing recommendations have expanded to include a broader scope of patients who may benefit from these therapies[6,16]. As a result, patients with advanced HER2-negative disease, triple-negative BC, and younger patients in general are now routinely tested[1]. Beyond germline mutations, somatic mutations can be identified through next-generation sequencing (NGS), both from tissue samples and liquid biopsies using ctDNA. Liquid biopsies provide a non-invasive, real-time approach to capture genomic alterations, allowing for dynamic monitoring of tumor evolution. Somatic testing has recently been upgraded in the latest ESMO guidelines update[17]. The upgrades include higher ESCAT scores for *ESR1* (IA), somatic
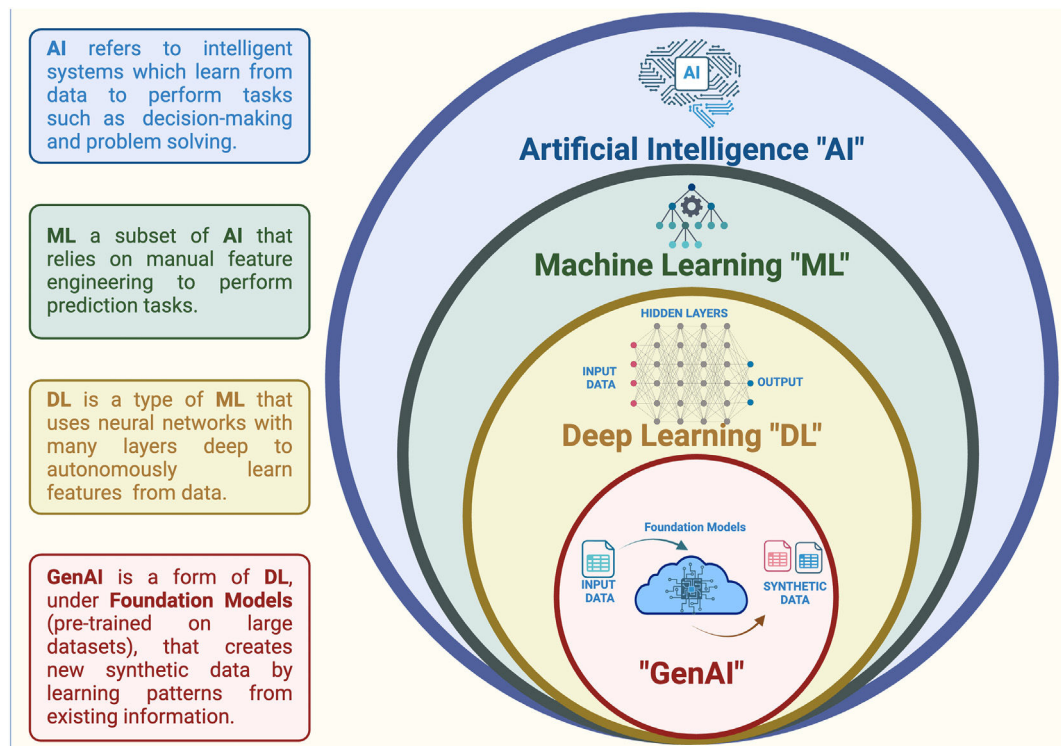
Figure 1 Overview of the hierarchical relationship between Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL), and Generative AI (GenAI).

BRCA mutations (IIB), *PTEN* (I/II), *AKT* (I/II), and germline pathogenic variants in *PALB2* (IIB), adding to the previously recommended germline *BRCA1/2* mutations (IA), *PIK3CA* (IA), *ERBB2* amplification (IA), and *ERBB2* hotspot mutations (IIB).

BC is not a monogenic disease, where a single altered gene drives the phenotype. Instead, multiple low-penetrant mutations act cumulatively[2]. It is a complex disease influenced not only by mutations but also by DNA variants that either increase or decrease the patient's risk. Single-nucleotide polymorphisms (SNPs), compared to driver mutations, have a significantly smaller effect size but are more common and overlooked. Conte et al. recently demonstrated how differences in SNPs in aromatase enzymes could determine sensitivity to endocrine therapy in breast cancer patients, underscoring the importance of larger population studies[19]. Moreover, genome-wide association studies (GWASs) enable the development of polygenic scores that can better refine risk assessment for future cancer risk and screening strategies as well as predict survival among cancer patients[18]. However, so far, its clinical use remains limited.

Genomic expression panels evaluating transcriptomics remain central for early BC patients and are routinely used to estimate prognosis and risk of recurrence. They can have a predictive role in selecting which patients will benefit most from specific treatments, such as chemotherapy. Panels such as the 21-gene recurrence score (Oncotype) or the PAM50 help guide adjuvant therapy decisions[20,21]. New RNA-based signatures have widened the scope by gaining traction on other subtypes, such as HER2DX for patients with HER2-positive disease and TNBCDX for triple-negative disease[22,23]. These are important advancements in genomic profiling, as several factors beyond the initial stage have been associated with patients' prognosis and/or treatment response[24]. However, decisions about escalation or de-escalation of systemic therapies are still based on traditional parameters, i.e., tumor size, nodal status, expression of the HR, and response to neoadjuvant therapy. This approach is insufficient to capture the full complexity of the disease. Therefore, a tool that integrates multiple variables is likely to outperform any single feature. To this end, the HER2DX assay was recently developed and validated, integrating multiple factors using ML algorithms. This is the first combined prognostic score based on clinicopathological and genomic variables in early-stage HER2 + BC, providing two independent scores to predict both long-term prognosis and the likelihood of pathological complete response (pCR) in patients with HER2 + early BC. HER2DX identifies a substantial proportion of patients with early-stage HER2+ BC who might not need additional therapies, such as pertuzumab, neratinib, or T-DM1, due to favorable survival outcomes with chemotherapy and trastuzumab alone. In addition, HER2DX can identify patients with high-risk disease who might need additional anti-HER2 therapies beyond trastuzumab. Similarly, TNBCDX is a genomic test for early-stage triple-negative BC, currently undergoing validation and standardization. Utilizing ML, it integrates clinical variables with tumor and immune-related factors to provide a comprehensive risk assessment that can guide treatment decisions. TNBCDX is trained on large datasets of patients with known outcomes, learning patterns that correlate with specific

clinical endpoints. In this case, ML incorporates genomic and clinical data into a predictive model that offers two key scores: a risk score and a pCR likelihood score.

RNA-based profiling in tumor tissues can identify complex biological processes grouped into predictive and prognostic subtypes[25]. However, tumor biopsies are invasive, challenging, and may not adequately reflect heterogeneity across metastatic sites. Liquid biopsies offer a minimally invasive, easily repeatable, and more comprehensive assessment of tumor heterogeneity by capturing genetic alterations from multiple tumor sites simultaneously and allowing for real-time monitoring of tumor evolution and treatment response. Circulating tumoral (ct)DNA serves as a non-invasive source of cancer DNA for analyzing tumor somatic genetic features. DNADX, a novel biomarker derived from ML analysis of multi-gene signatures in ctDNA, can effectively capture relevant biological and complex phenotypic features like those identified through tumor tissue DNA or RNA profiling. These features include 1) tumor cell proliferation, 2) activation of the ER pathway, 3) retinoblastoma loss-of-heterozygosity (RB-LOH) status, 4) TP53 activation status, and 5) the newly identified DNA-based intrinsic subtypes, among others[26]. Recent work has demonstrated that ctDNA-based genomic signature tracking RB-LOH and DNA-based subtypes can predict poor outcomes following ET combined with CDK4/6i. This is consistent across different plasma tumor fractions and has been validated in two independent cohorts[26,27]. HER2DX, TNBCDX, and DNADX exemplify how AI can integrate clinical and genomic data to optimize patient care.

While each of these approaches provides valuable insights and has improved treatment personalization, they offer only a partial view of tumor biology and are limited by their reliance on predefined features. To fully capture the heterogeneity of breast cancer and personalize treatment approaches, a more integrative strategy, using, for instance, DL, could make a difference by offering the ability to analyze vast and unstructured data to identify previously hidden patterns.

Deep learning

New AI technologies have enabled the transition from traditional "shallow" ML models to "deep" learning (Fig. 1). DL mimics how the human brain works, recognizing patterns and making decisions without requiring manually identified features. It uses neural networks, many layers ("deep") of interconnections that process and analyze data. Each layer examines the data, starting from simple features and working its way up to more complex patterns. DL models are trained on large amounts of data, allowing them to improve over time in predicting outcomes from unseen and unstructured data without the need to pre-define what is important. This makes DL a particularly promising tool for biomedical data analysis, driven by the rapid growth of multi-omics[28].

Since the first description of estrogen receptors (ER) in the 1950s, and later published in the 1970s, followed by the identification of HER2 overexpressing cells in the 1980s, the use of these and other biomarkers, such as Ki67, which helps determine cell proliferation rate, has been central to better

define an extremely heterogeneous disease. Dividing BC into four distinct subtypes was the first step toward personalizing treatment. Yet, with the advent of new therapies and broader molecular techniques, this early description seems now insufficient to thoroughly describe the diverse behavior of tumors.

Until recently, HER2 status was classified as either positive or negative, based on IHC and in situ hybridization (ISH), following the American Society of Clinical Oncology/College of American Pathologists (ASCO/CAP) guidelines[29]. However, with the development of novel drugs, such as trastuzumab-deruxtecan, an antibody-drug conjugate that targets HER2 receptors, a new category, "HER2-low", has emerged[30]. This has led to new definitions and criteria that are not easily reproducible and standardized. A recent study by Fernandez et al. showed an 11% discordance between pathologists in distinguishing HER2 3+ from non-3+, and up to 41% when differentiating HER2 1+ or HER2 2+ from HER2 0. Differences have significant clinical implications[31]. Consequently, many groups are working on developing AI-assisted models to improve the accuracy, reliability, and concordance of HER2 interpretation across pathologists. Palm C et al. examined the effectiveness of AI-assisted workflows in evaluating HER2, using IHC and ISH, compared to the traditional pathologist asse[32]. Their study showed that using AI/DL to combine pathologist-based assessments of IHC and ISH is feasible, resulting in a Cohen's κ of 0.94 under ASCO/CAP recommendations. This enables AI/DL, both in primary and metastatic BC, to recognize HER2-low group tumors. Similarly, Jaber et al. (BCR 2020) presented a DL model that accurately classified the five molecular subtypes - luminal A/B, HER2-enriched, basal-like, and normal-like - using histopathology.

Ki67 is a well-known predictive and prognostic biomarker and holds value as a dynamic biomarker of response to treatment[33]. Still, the lack of standardized measurement protocols further limits its true clinical utility. Some studies have demonstrated the value of automated imaged-based methods for quantifying Ki67, rather than relying on visual scoring. Rimm et al. conducted an international trial with an automated scoring system that demonstrated high reproducibility, showing less variability than manual assessments and the potential to enhance standardization[34]. When evaluating Ki67, intra-tumor heterogeneity may pose a challenge. Digital image analysis (DIA) allows the automated visualization of Ki67 heterogeneity and the detection of Ki67 hotspots, as demonstrated by Plancoulaine et al., who developed a DIA-generated system based on the hexagonal tiling of digital image analysis data, offering more precise prognostic information and improving the clinical value of Ki67 assessments[35].

Regarding histologic/phenotypic entities, new strategies seek to combine genomic and pathological characteristics to group patients. Recent work using a DL model (convolutional neural networks), applied to whole-slide imaging utilizing *CDH1* bi-allelic mutations rather than histology as the gold standard, proposed a proof-of-concept diagnostic approach for lobular BC. The most interesting finding was that those cases that were initially defined by the model as false positive, as no bi-allelic mutations of *CDH1* were detected, were truly positive with novel *CDH1* inactivating genetic alterations initially overlooked. Underscoring how the DL

model can help recognize patterns that the human eye might easily miss[36].

Beyond tumor cell-specific biomarkers, the tumor microenvironment (TME) has become central in understanding tumoral heterogeneity. Amgad et al. recently presented the Histomic Prognostic Signature (HiPS), a comprehensive interpretable scoring system designed to determine survival risk based on breast TME morphology[37]. HiPS uses DL to map cellular and tissue structures, measuring epithelial, stromal, immune, and spatial interaction features. Driven mainly by immune and stromal features, HiPS outperformed pathologists in predicting survival risk among patients, regardless of stage and clinical biomarkers. The ENLIGHT-DeepPT model was recently introduced as a DL framework that accurately predicts gene expression from histopathology images, using inferred expression values that offer a predictive tool for targeted and immune therapies[10].

AI/DL has the potential to automate and standardize widely available biomarkers, reducing variability and increasing reliability and reproducibility. Its ability to interpret and recognize complex patterns within the TME can lead to comprehensive models that may offer personalized treatment plans.

Integrating multi-OMICS with AI

Building on DL's ability to uncover complex patterns, AI plays a pivotal role in integrating multi-omics data, providing a comprehensive view of tumor biology and its evolution. By combining genomic data, such as DNA repair mutations or APOBEC signatures, with gene expression and tumor microenvironment (TME) features, AI models generate predictive insights without requiring manual prioritization of features.

Several AI/DL models have shown promising results in delivering personalized and precise predictions for BC outcomes. DeepMO, which uses the TCGA dataset, integrates multi-omics layers to predict patient prognosis and response to therapy. Similarly, DeepGene helps better classify cancer types using somatic point mutations[38,39]. However, heterogeneity in data collection and lack of standardization across different platforms are key limitations to wider clinical adoption. To overcome these challenges, global initiatives have been launched to develop standardized and accurate AI/DL platforms. For instance, the OPTIMA (Optimal Treatment for Patients with Solid Tumors in Europe Through Artificial Intelligence) is a public-private research program that aims to improve patient care using AI working to establish secure and large-scale data platforms, develop advanced analytics and AI models to improve clinical guidelines and create AI-based decision support tools.

Future directions

While these new technologies and expanding treatment strategies hold great promise, challenges remain. First, AI depends on data, consequently, the standardization of data, which ensures high-quality and complete records, is essential. DL models are only as good as the data provided; therefore, datasets must also be diverse to avoid perpetuating biases. Second, ethical issues must be addressed, as ensuring safety and privacy are paramount. Accountability must be regulated as AI-driven clinical recommendations become more prevalent. Third, clinicians must learn to trust and understand AI-driven information. Although several trials using gene expression platforms have shown that many patients can de-escalate, the medical community can be slow to adopt changes.

Generative AI (GenAI), an application of foundation models, a more recent subset of AI based on large pre-trained models, can create new data, such as text, images, or synthetic datasets, by learning patterns from existing information. By generating synthetic data or augmenting existing datasets, GenAI could help develop more robust models and overcome some of these limitations, such as data heterogeneity and quality. Additionally, autonomous AI can plan, execute, and optimize multi-step workflows, reducing the need for constant human intervention and addressing the current limitations of single-purpose AI models by integrating multimodal data and enhancing problem-solving efficiency. However, rigorous evaluation and human oversight remain essential to safely implement these systems in oncology.

Conclusions

Advancements in oncology come with a growing economic burden, placing significant pressure on our resource-limited health systems. Therefore, with more multi-omics data on cancer, tailoring treatment strategies is mandatory. Integrating genomics and transcriptomics enhances our understanding of tumor biology, reduces overtreatment, and optimizes resource use, paving the way for personalized, cost-effective care.

AI in cancer care is evolving from simpler ML models that rely on predefined features to more advanced DL models capable of identifying patterns from large, unstructured datasets. As AI models evolve and incorporate a wider scope of information, their success and clinical impact will depend on robust data infrastructure, collaboration, and strong ethical governance.

Funding

Ethical considerations

This review synthesizes data from previously published studies and did not involve direct interaction with human subjects. Consequently, specific ethical approval was not required.

Author contributions statement

RGB, BW, and ES wrote the manuscript, and all authors planned and revised the manuscript.

Conflict of interests

The authors declare no conflicts of interest related to this work.

References

1. Bedrosian I, Somerfield MR, Achatz MI, Boughey JC, Curigliano G, Friedman S, et al. Germline testing in patients with breast cancer: ASCO-Society of Surgical Oncology Guideline. J Clin Oncol. 2024;42(5):584−604.

2. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primers. 2019;5 (1):66.

3. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747−52.

4. Prat A, Carey LA, Adamo B, Vidal M, Tabernero J, Cortés J, et al. Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer. JNCI: J Natl Cancer Inst. 2014;106(8).

5. Bottosso M, Mosele F, Michiels S, Cournède PH, Dogan S, Labaki C, et al. Moving toward precision medicine to predict drug sensitivity in patients with metastatic breast cancer. ESMO Open. 2024;9(3), 102247.

6. Tutt ANJ, Garber JE, Kaufman B, Viale G, Fumagalli D, Rastogi P, et al. Adjuvant olaparib for patients with BRCA1- or BRCA2-mutated breast cancer. N Engl J Med. 2021;384(25):2394−405.

7. Turner NC, Oliveira M, Howell SJ, Dalenc F, Cortes J, Gomez Moreno HL, et al. Capivasertib in hormone receptor−positive advanced breast cancer. N Engl J Med. 2023;388(22):2058−70.

8. André F, Ciruelos E, Rubovszky G, Campone M, Loibl S, Rugo HS, et al. Alpelisib for PIK3CA -mutated, hormone receptor−positive advanced breast cancer. N Engl J Med. 2019;380(20):1929−40.

9. Bidard FC, Kaklamani VG, Neven P, Streich G, Montero AJ, Forget F, et al. Elacestrant (oral selective estrogen receptor degrader) versus standard endocrine therapy for estrogen receptor−positive, human epidermal growth factor receptor 2−negative advanced breast cancer: results from the randomized phase III EMERALD trial. J Clin Oncol. 2022;40(28): 3246−56.

10. Hoang DT, Dinstag G, Shulman ED, Hermida LC, Ben-Zvi DS, Elis E, et al. A deep-learning framework to predict cancer treatment response from histopathology images through imputed transcriptomics. Nat Cancer. 2024;5(9):1305−17.

11. Piccart M, van't Veer LJ, Poncet C, JMN Lopes Cardozo, Delaloge S, Pierga JY, et al. 70- gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomized MINDACT trial with an exploratory analysis by age. Lancet Oncol. 2021;22(4):476−88.

12. Coates AS, Winer EP, Goldhirsch A, Gelber RD, Gnant M, Piccart-Gebhart M, et al. Tailoring therapies−improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer. Ann Oncol. 2015;26(8):1533−46.

13. Heredia A, Walbaum B, Vidal M, Itriago L, Camus M, Dominguez F, et al. Suboptimal use of ovarian function suppression in very young women with early breast cancer: a real-world data study. Breast Cancer Res Treat. 2024;203(1):173−9.

14. Zhang B, Shi H, Wang H. Machine learning and AI in cancer prognosis, prediction, and treatment selection: a critical approach. J Multidiscip Healthc. 2023;16:1779−91.

15. Lee M. Deep learning techniques with genomic data in cancer prognosis: a comprehensive review of the 2021-2023 Literature. Biology (Basel). 2023;12(7).

16. Litton JK, Rugo HS, Ettl J, Hurvitz SA, Gonçalves A, Lee KH, et al. Talazoparib in patients with advanced breast cancer and a germline BRCA mutation. N Engl J Med. 2018;379(8):753−63.

17. Mosele MF, Westphalen CB, Stenzinger A, Barlesi F, Bayle A, Bièche I, et al. Recommendations for the use of next-generation sequencing (NGS) for patients with advanced cancer in 2024: a report from the ESMO Precision Medicine Working Group. Ann Oncol. 2024;35(7):588−606.

18. Yang X, Kar S, Antoniou AC, Pharoah PDP. Polygenic scores in cancer. Nat Rev Cancer. 2023;23(9):619−30.

19. Conte B, Boni L, Bisagni G, Durando A, Sanna G, Gori S, et al. SNP of aromatase predict long-term survival and aromatase inhibitor toxicity in patients with early breast cancer: a biomarker analysis of the GIM4 and GIM5 trials. Clin Cancer Res. 2023;29(24):5217−26.

20. Sparano JA, Gray RJ, Makower DF, Pritchard KI, Albain KS, Hayes DF, et al. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. N Engl J Med. 2018;379(2): 111−21.

21. Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the PAM50 Risk of Recurrence score in 1478 postmenopausal patients of the ABCSG- 8 trial treated with adjuvant endocrine therapy alone. Ann Oncol. 2014;25(2):339−45.

22. Prat A, Guarneri V, Pascual T, Brasó-Maristany F, Sanfeliu E, Paré L, et al. Development and validation of the new HER2DX assay for predicting pathological response and survival outcome in early-stage HER2-positive breast cancer. EBioMedicine. 2022;75, 103801.

23. Martín M, Stecklein SR, Gluz O, Villacampa G, Monte-Millán M, Nitz U, et al. TNBC-DX genomic test in early-stage triple-negative breast cancer treated with neoadjuvant taxane- based therapy. Annals of Oncology. 2024;36(2):158−71.

24. Leon-Ferre RA, Jonas SF, Salgado R, Loi S, de Jong V, Carter JM, et al. Tumor-infiltrating lymphocytes in triple-negative breast cancer. JAMA. 2024;331(13):1135.

25. Prat A, Chaudhury A, Solovieff N, Paré L, Martinez D, Chic N, et al. Correlative biomarkeranalysis of intrinsic subtypes and efficacy across the MONALEESA phase III studies. J Clin Oncol. 2021;39(13):1458−67.

26. Prat A, Brasó-Maristany F, Martínez-Sáez O, Sanfeliu E, Xia Y, Bellet M, et al. Circulating tumor DNA reveals complex biological features with clinical relevance in metastatic breast cancer. Nat Commun. 2023;14(1):1157.

27. Brasó-Maristany F, Cortés J, Pérez-García JM, Vega-León R, Paré L, Villacampa G, et al. Abstract PO2−15-06: ctDNA-based DNADX in hormone receptor-positive and HER2- negative (HR+/HER2-) advanced breast cancer following endocrine therapy and CDK4/6 inhibition: a correlative analysis from the randomized phase 2 PARSIFAL trial. Cancer Res. 2024;84(9_Supplement), PO2-15-06.

28. Ma B, Meng F, Yan G, Yan H, Chai B, Song F. Diagnostic classification of cancers using extreme gradient boosting algorithm and multi-omics data. Comput Biol Med. 2020;121, 103761.

29. Wolff AC, Hammond MEH, Allison KH, Harvey BE, Mangu PB, Bartlett JMS, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Focused Update. J Clin Oncol. 2018;36(20):2105−22.

30. Modi S, Jacot W, Yamashita T, Sohn J, Vidal M, Tokunaga E, et al. Trastuzumab deruxtecan in previously treated HER2-low advanced breast cancer. N Engl J Med. 2022;387(1):9−20.

31. Fernandez AI, Liu M, Bellizzi A, Brock J, Fadare O, Hanley K, et al. Examination of low ERBB2 protein expression in breast cancer tissue. JAMA Oncol. 2022;8(4):1−4.

32. Palm C, Connolly CE, Masser R, Padberg Sgier B, Karamitopoulou E, Simon Q, et al. Determining HER2 status by artificial intelligence: an investigation of primary, metastatic, and HER2 low breast tumors. Diagnostics (Basel). 2023;13(1).

33. Criscitiello C, Disalvatore D, De Laurentiis M, Gelao L, Fumagalli L, Locatelli M, et al. High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in luminal B HER2 negative and node-positive breast cancer. Breast. 2014;23(1):69−75.

34. Rimm DL, Leung SCY, McShane LM, Bai Y, Bane AL, Bartlett JMS, et al. An international multicenter study to evaluate reproducibility of automated scoring for assessment of Ki67 inbreast cancer. Mod Pathol. 2019;32(1):59−69.

35. Plancoulaine B, Laurinaviciene A, Herlin P, et al. A methodology for comprehensive breast cancer Ki67 labeling index with intra-tumor heterogeneity appraisal based on hexagonal tiling of digital image analysis data. Virchows Arch. 2015;467:711−22.

36. Pareja F, Dopeso H, Wang YK, Gazzo AM, Brown DN, Banerjee M, et al. A genomics-driven artificial intelligence − based model classifies breast invasive lobular carcinoma and discovers *CDH1* inactivating mechanisms. Cancer Res. 2024;84(20):3478−89.

37. Amgad M, Hodge JM, Elsebaie MAT, Bodelon C, Puvanesarajah S, Gutman DA, et al. A population-level digital histologic biomarker for enhanced prognosis of invasive breast cancer. Nat Med. 2024;30(1):85−97.

38. Lin Y, Zhang W, Cao H, Li G, Du W. Classifying breast cancer subtypes using deep neural networks based on multi-omics data. Genes (Basel). 2020;11(8).

39. Yuan Y, Shi Y, Li C, Kim J, Cai W, Han Z, et al. DeepGene: an advanced cancer type classifier based on deep learning and somatic point mutations. BMC Bioinform. 2016;17(Suppl 17): 476.