Advancement of prognostic models in breast cancer: a narrative review

Ningning Min[1,2], Yufan Wei[1,2], Yiqiong Zheng[2], Xiru Li[2]

[1]School of Medicine, Nankai University, Tianjin, China; [2]Department of General Surgery, Chinese People's Liberation Army General Hospital, Beijing, China

Contributions: (I) Conception and design: X Li, N Min; (II) Administrative support: X Li; (III) Provision of study materials or patients: X Li; (IV) Collection and assembly of data: N Min; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Xiru Li. Department of General Surgery, Chinese People's Liberation Army General Hospital, Beijing 100853, China. Email: 2468li@sina.com.

Objective: To provide a reference for clinical work and guide the decision-making of healthcare providers and end-users, we systematically reviewed the development, validation and classification of classical prognostic models for breast cancer.

Background: Patients suffering from breast cancer have different prognosis for its high heterogeneity. Accurate prognosis prediction and risk stratification for breast cancer are crucial for individualized treatment. There is a lack of systematic summary of breast cancer prognostic models.

Methods: We conducted a PubMed search with keywords "breast neoplasm", "prognostic model", "recurrence" and "metastasis", and screened the retrieved publications at three levels: title, abstract and full text. We identified the articles presented the development and/or validation of models based on clinicopathological factors, genomics, and machine learning (ML) methods to predict survival and/or benefits of adjuvant therapy in female breast cancer patients.

Conclusions: Combining prognostic-related variables with long-term clinical outcomes, researchers have developed a series of prognostic models based on clinicopathological parameters, genomic assays, and medical figures. The discrimination, calibration, overall performance, and clinical usefulness were validated by internal and/or external verifications. Clinicopathological models integrated the clinical parameters, including tumor size, histological grade, lymph node status, hormone receptor status to provide prognostic information for patients and doctors. Gene-expression assays deeply revealed the molecular heterogeneity of breast cancer, some of which have been cited by AJCC and National Comprehensive Cancer Network (NCCN) guidelines. In addition, the models based on the ML methods provided more detailed information for prognosis prediction by increasing the data dimension. Combined models incorporating clinical variables and genomics information are still required to be developed as the focus of further researches.

Keywords: Breast cancer; prognostic model; clinicopathological model; gene expression assay; machine learning model (ML model)

Introduction

Breast cancer is one of the most common malignant tumors in women, which is highly heterogeneous in molecular changes, cellular composition, therapeutic response, and clinical outcomes. Based on the transcriptional level, breast cancer can be divided into five intrinsic subtypes, luminal A (LumA), luminal B (LumB), HER2-enriched, basal-like, and normal-like (1). According to immunohistochemistry

staining, breast cancer can be further categorized as luminal subtypes (A/B), HER2-overexpression subtype and triple negative breast cancer (TNBC). With the number of available treatment options, it becomes increasingly difficult and complex to make the most appropriate treatment choice for different breast cancer subtypes. Therefore, more personalized management of breast cancer are required.

A variety of clinicopathologic factors, including tumor size, lymph node status, and histological grade, together with molecular markers, such as estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), and Ki-67 were used to describe the specific features and prognosis of breast cancer (2). However, the traditional clinicopathological factors cannot accurately recognize low-risk groups. The toxic effects and high costs of overtreatment remain the critical issues in the treatment of low-risk patients. Meanwhile, the high-risk patients confront the risk of undertreatment. Therefore, the focus of current individualized treatment for breast cancer patients is to predict the prognosis accurately and distinguish high-risk group from the majority of patients with low recurrence risk who may avoid chemotherapy.

Prognostic model refers to the use of statistical methods to determine the quantitative relationship between the risk factors and the probability of clinical outcomes based on the patient's disease state. Breast cancer prognostic models can help clinicians and healthcare providers make more informed medical decisions on chemotherapy exemption. Those models based on clinical variables, such as Nottingham Prognostic Index (NPI), Adjuvant! Online and PREDICT, are the earliest and most widely used models. With the more profound comprehension of breast biology, a series of updates made since launch have introduced new prognostic variables (HER2, Ki67) and refined the prediction algorithm. However, conventional clinicopathological models cannot reveal the extensive molecular heterogeneity seen between breast tumors even in the hormone-receptor-positive subgroup, leading to bias in predicting clinical outcomes. To achieve a more accurate prediction of the prognosis, a series of prognostic models added genomic information on the basis of clinicopathological factors and molecular markers. The AJCC tumor staging system recommends the use of gene expression assays for prognostic stratification, and the 21-gene assay (21-GA) Oncotype DX® was recommended as class I evidence (2). The ASCO biomarker guidelines also approved the use of 70-gene test MammaPrint® to provide information on the benefits of chemotherapy

for ER-positive patients with high-risk clinical features. MammaPrint also became the only test approved for patients with 1–3 positive lymph nodes (3). In recent years, artificial intelligence technology has developed rapidly, and data-driven machine learning (ML) methods have been developed to realize the prediction of breast cancer prognosis by learning statistical rules or patterns from historical data. The use of ML methods to analyze the texture information of histopathological sections and imaging images broke through the limitations of human eye recognition by improving the data dimension, and provided more detailed and reliable information for the prognosis prediction. In this study, the development, validation, and classification of the existing classical prognostic models of breast cancer were reviewed to provide a reference for clinical work.

We present the following article in accordance with the Narrative Review checklist (available at https://dx.doi.org/10.21037/gs-21-441).

Model development

Variable screening and variable processing

Breast cancer is a highly heterogeneous disease, and the prognosis of patients is affected by a variety of factors. The selection of prognostic variables is the basis for the construction of a prognosis prediction model. For traditional clinical variables, statistically significant prognostic factors were identified preliminarily through performing univariate and multivariate analyses by t-test, Logistic regression, Cox proportional risk regression, Lasso regression, etc. Then, candidate variables were further screened by adjusting the influence of confounders and combining them with clinical practice experience. Finally, the parameters are transformed in the form necessary to enter the prediction model. Among the clinicopathological factors, lymph node status, tumor size, tumor grade, age at diagnosis, and ER status were most commonly used to develop prognostic models for breast cancer (4).

For gene expression assays, these statistical methods mentioned are also used to select genes related to the prognosis of breast cancer from published literature, genome databases, or DNA array experiments based on tumor tissue. In addition, cluster analysis is an effective method to obtain genomes with similar functions or characteristics. Then, using minimization algorithms such as top "N" t-test statistics, top cluster index scores and,

variance inflation factor (VIF), to remove redundant and irrelevant parts, researchers will obtain the minimized gene sets with prognostic value.

Model construction

Prognostic models were trained by combining prognostic factors with the long-term clinical outcomes in the development cohorts. Mathematical prediction models are divided into parametric models, semi-parametric models, and non-parametric models. Parameterized models include Linear regression, Logistic regression, Poisson regression, and Discriminant analysis. Cox proportional hazards regression and Competitive hazards model are the main semi-parametric models. Non-parametric models, namely ML algorithms, mainly include K-Nearest Neighbor (KNN) algorithm, support vector machines (SVM), penalty regression model, artificial neural networks, Bayesian networks, random forest, decision trees and so on. Among them, Cox proportional hazards regression was the most commonly used method for model development, followed by artificial neural networks, decision trees, Logistic regression, and Bayesian networks (4). Following parameter estimation and model fitting, these prognostic models can be presented as formulas, nomograms, web calculators, or scoring systems to quantify the likelihood or risk of the clinical outcome and classify patients into different risk groups.

Model validation

The evaluation of the prognostic model includes internal and external verification. Internal verification refers to assessing the reproducibility of the model in the same patient cohort as the development cohorts, or the development cohorts with longer follow-up, or specific subgroups of the development cohorts, or the combination of the development cohorts and newly recruited patients in the same centers. The sampling methods for internal validation mainly include cross-validation, random-splitting, bootstrap and so on. External verification refers to validating the model's generalizability in patients independent of the development cohort, comparing the predicted outcomes with observations, predictions from other models, or predictions from single prognostic factors.

The evaluation of the prognostic model mainly includes four aspects. (I) Discrimination, the extent to which the model distinguishes patients with the outcomes from those without the outcomes (disease/no disease, effective/ineffective, death/survival). Harrell's C-index/AUC [area under the receiver operating characteristic (ROC) curve], precision-recall curve, Kaplan-Meier curve, Log-rank test, net reclassification improvement (NRI), and integrated discrimination improvement (IDI) are commonly used metrics to evaluate the discrimination of different models. (II) Calibration, the level of agreement between the predicted and observed outcomes, can be measured by the calibration plot and Hosmer-Lemeshow goodness-of-fit test. (III) Overall performance, the evaluation integrating the discrimination and calibration of the model. Explained variation R2, Brier scores, Akaike information criterion (AIC), Bayesian information criterion (BIC) were commonly used statistical indexes. (IV) Clinical usefulness, the ability to provide clinicians with decision-making guidance, can be evaluated based on accuracy rate, sensitivity/specificity, Kaplan-Meier survival curve, and decision curve analysis (DCA). In addition, some studies comparing two or more models used the Kappa coefficient (κ) and correlation coefficient (Pearson or Spearman) to test for consistency between the models.

Model classification

A PubMed search using the search terms "breast neoplasm", "prognostic model", "recurrence", and "metastasis" was conducted, regardless of the type of article. The retrieved publications were screened in three levels-titles, abstracts, and full texts. From each selected article, relevant information was extracted to form the framework of this article. We classified prognostic models into clinicopathological models, gene-expression assays, and ML models, and performed a secondary search based on the representative model of each section. The prediction accuracy and clinical utility of these models were demonstrated.

Clinicopathological model

NPI

NPI is a prognostic index for patients with early-stage primary breast cancer to predict survival based on lymph node stage, tumor size, and histological grade. In 1982, Haybittle *et al.* (5) attempted to combine multiple prognostic factors into a prognostic index called NPI, using the Cox proportional hazard model in a retrospective multivariate analysis. The NPI allows the stratification of patients into three different prognostic groups and predicts the possible

2818

Min et al. Advancement of prognostic models in breast cancer

benefit of patients receiving adjuvant therapy. The use of the NPI has been subsequently validated in several prospective cohorts (6,7), and extensive multicenter studies independently (8,9). To provide a more accurate survival estimate after breast cancer surgery, Blamey *et al.* (10) divided the patient cohorts into ten different prognostic groups based on the NPI value and developed interpolation models using continuous NPI, NPI-squared, and NPI-cubed covariates to evaluate the survival of individual patients. Those interpolation models using higher-resolution show better performance than the six-group NPI model initially described by the research group (11). In addition, NPI retains the ability to stratify and predict survival in patients with TNBC, even though TNBC seems to disseminate to axillary nodes and bones less frequently than non-TNBC, presenting a preferential hematogenous route with a propensity to develop metastatic deposits in viscera (12,13). Besides, the modified NPI (MNPI), obtained by adding the modified Scarff-Bloom-Richardson (SBR) grade rather than the SBR grade to evaluate the histological grade, could be used to stratify patients with stage I to III TNBC according to prognosis (14).

Van Belle *et al.* (15) observed that PR and HER2 provided independent prognostic information and could be added to the NPI to develop a new prognostic index, iNPI, for improved disease-free survival (DFS) prediction. The iNPI demonstrated good discrimination and classification accuracy in two independent cohorts, improving the predictive evaluation and risk stratification in patients with operable breast cancer. In addition, Winzer *et al.* (16) obtained an extended NPI with improved prognostic ability by adding the information from hormonal status and using the complete information from the three NPI components, which were extracted through multi-variable fractional polynomials and further modern statistical methods.

Considering that molecular features of breast cancer are the critical driver of tumor behavior, Rakha *et al.* (17) developed a two-tier prognostic scoring system, NPI +, which combined biomarkers and traditional clinicopathologic variables. In NPI +, the initial assessment determined seven core biological classes of the tumor based on ten breast cancer-related biomarkers, and a second-level analysis identified six clinicopathological prognostic factors associated with breast cancer-specific survival (BCSS) subsequently, resulting in tailored NPI-like formulae for each biological class. NPI + not only overcomes the variability in the prognostic power of each individual clinicopathologic factor in the different molecular classes to achieve more accurate

stratification of survival for breast cancer patients but also provides more complex personalized treatment tools for breast cancer patients. As a predictor of prognosis, the reproducibility and effectiveness of NPI + have been verified in multiple independent European coalitions of breast cancer patients (18,19).

Adjuvant! Online

Adjuvant! Online is a web-based tool that predicts survival and benefits of adjuvant therapy for women with early-stage breast cancer (stage I–III). After inputting patient information (age, menopausal status, comorbidity estimate) and tumor staging and characteristics (tumor size, number of positive axillary nodes, ER status), the model used actuarial analysis to predict the outcomes of patients with or without adjuvant therapy (endocrine therapy and/or chemotherapy). Part of the forecast results is presented in both numerical and graphical formats. For Adjuvant! Online, the 10-year survival estimates for patients without adjuvant therapy, including overall survival (OS), BCSS, and event-free survival (EFS), were derived from women aged 35 to 69 years old who were diagnosed between 1988 and 1992 in the United States and recorded in the Surveillance, Epidemiology and End Results (SEER) registry. Estimates of the efficacy of adjuvant therapy were derived from the Early Breast Cancer Trialists' Collaborative Group 1998 meta-analysis data (20).

In contrast with the other inputs, for which Adjuvant! Online used widely accepted definitions to minimize errors due to the subjective nature of user interpretation or judgment, comorbidity estimate lacks a clinically implemented scoring system. Ozanne *et al.* (21) confirmed that the input of comorbidity estimate affected the predictions of Adjuvant! Online significantly, especially for women aged 60 years and above, through single-variable deterministic sensitivity analysis. In addition, because Adjuvant! was developed in a cohort with relatively young patients, and the additional comorbidities of elderly patients led to an increased risk of competing deaths, the predictive performance of Adjuvant! Online in the elderly patient population was questioned widely. Applying two comorbidity classification models in the population-based FOCUS cohort of elderly patients respectively (22), Adjuvant! showed poor calibration for the 10-year OS and recurrence rate. Therefore, a prediction method specific to elderly breast cancer patients is needed. Tumor characteristics as well as detailed and standardized patient information should be included in the model to take into account the considerable

heterogeneity of the older patients with breast cancer.

Adjuvant! has been updated several times to include long-term follow-up data, more reliable cause-of-death information, and the latest evidence of therapeutic efficacy, since development (23-25). In the patient cohort from Canadian Breast Cancer Outcomes Unit (BCOU) database (23), Adjuvant! Online predicted overall 10-year OS, BCSS, and EFS accurately, which were significantly overestimated in the subgroup of patients under 35 years of age and positive lymphatic vascular invasion (LVI). A manual adjustment using the prognostic factor influence calculator (PFIC) reduced the apparent discrepancy, but predictions by Adjuvant! Online were still overly optimistic for BCSS and EFS. Similar results were obtained in the Netherlandish (24) and French (26) studies, in which the OS and BCSS of patients under 40 were overestimated, and the BCSS of ER-negative patients was underestimated, although the model showed good calibration and discrimination performance. The estimated bias in these subgroups of patients suggests consideration of including youth, LVI, and HER2 status in the new version of Adjuvant! Online algorithm (23,24). Hajage *et al.* (26) found that HER2 status, mitochondrial index (MI), and Ki67 added independent predictive information to Adjuvant! as well. Notably, in the cohorts from the United Kingdom (25), Korea (27), Taiwan (28), and Malaysia (29), the OS/BCSS/EFS was significantly overestimated by Adjuvant! Online, and OS and BCSS in almost all subgroups were overestimated. The differences in results from different regions might be correlated with differences in sample size, patient distribution, and ethnicity.

PREDICT

PREDICT is an online prognostication and treatment benefit tool developed in the United Kingdom, which was based on the survival data recorded by the Eastern Cancer Registration and Information Centre (ECRIC) for 5,694 women diagnosed between 1999 and 2003 (30). In PREDICT, breast cancer specific mortality and mortality from other causes (competing mortality) were modeled separately, while ER-negative and ER-positive tumors were modeled separately too. Information obtained from ECRIC included age at diagnosis, lymph node status, tumor size, histological grade, ER status, mode of detection, information on local therapy, and type of adjuvant systemic therapy. However, age at diagnosis was not found to be significantly associated with breast cancer specific mortality and was excluded from subsequent models (PREDICT v1.1). Although tending to overpredict mortality, PREDICT v1.1 showed good calibration and discrimination performance in both the ECRIC and the WMCIU dataset for survival and treatment benefits, except that the fit was less good for some subgroups in WMCIU. PREDICT v1.1 has been validated in patient cohorts from British Columbia and performed well across all age subgroups except in women aged 20–35 where it under-predicted the actual mortality (31).

PREDICT v1.2, known as PREDICT +, was developed after introducing HER2 status as a prognostic marker into PREDICT v1.1, which incorporated an external estimate of the hazard ratio (HR) associated with HER2 positivity. PREDICT v1.2 provided better calibration performance for all-cause mortality and breast cancer specific mortality, especially in the specific mortality of HER2-positive patient subgroups. At the same time, PREDICT v1.2 allowed the calculation of trastuzumab treatment benefit estimates as well (32).

Similarly, PREDICT v1.3 was generated by applying the HR associated with the tumor proliferative marker Ki-67 to the baseline hazards used in PREDICT v1.2, which accommodates the prognostic differences between LumA and LumB subtypes in ER-positive tumors. PREDICT v1.3 showed better calibration (goodness-of-fit, P=0.065) and discrimination (AUC =0.7611 *vs.* 0.7676, P=0.005) performance for breast cancer specific mortality in ER-positive tumors (33). In addition, Drukker *et al.* (34) attempt to add the 70-gene signature (70-GS) (MammaPrint™) into PREDICT v1.2 in the patients with early-stage breast cancer and found no significant improvement in 5- or 10-year survival predictions.

Despite two modifications, PREDICT v1.3 still underestimated breast cancer-specific mortality in women diagnosed at 40 years old or younger, particularly those with ER-positive disease. Besides, PREDICT v1.3 used discrete categories for tumor size and node status, resulting in 'step' changes in risk estimations on moving from one category to the next. Therefore, PREDICT v2 re-included the age at diagnosis into the prediction of breast-cancer specific death, and re-recorded tumor size and node status as continuous variables, significantly increasing the predictive value of breast-cancer specific death risk in ER-positive young patients (35). The calibrated PREDICT v2 demonstrated higher accuracy in outcome prediction and greater population-generality across a large sample of the Scottish SCR dataset (36).

It is worth noting that using different basic models for

2820

Min et al. Advancement of prognostic models in breast cancer

ER-positive and ER-negative tumors is a crucial feature of PREDICT. In the PREDICT model, the effects of other prognostic variables vary with ER status, particularly in HER2 status. The HR for HER2-positive disease decreases over time in women with ER-positive breast cancer, while the effect of HER2 in women with ER-negative breast cancer is not time-dependent (32). Another feature of PREDICT is that the detection mode (screening or clinical) is used as one of the model input parameters, taking into account the additional survival advantage of breast screening.

Special clinicopathologic prognostic models

Bremer *et al.* (37) developed a biologic signature named DCISionRT for the calculation of individual decision score (DS), which combined molecular markers and clinicopathological factors associated with recurrence or progression of ductal carcinoma in situ (DCIS) in a nonlinear model. In the cohort of DCIS patients with or without radiotherapy after breast-conserving surgery (BCS), the elevated group of DS showed higher risk for 10-year invasive breast cancer or any ipsilateral breast events, and received significant radiotherapy benefit.

Qiu *et al.* (38) found that four groups of breast cancer stem cells (BCSCs) (ALDH1A3, CD44$^+$/CD24$^-$, ITGA6, and PROCR) were identified as correlated with the relapse-free survival (RFS) in early-stage breast invasive ductal carcinoma (BIDC). The correlated biomarkers were integrated into a prognostic panel to calculated relapse risk score (RRS). RRS demonstrated significant predictive power for 5-year recurrence rates, independent of age at diagnosis and tumor size. Notably, there was no difference for RFS between those hormone-treated patients and non-treated patients in the high-risk score group (P=0.860), indicating that high RRS score ER-positive patients may not benefit from traditional endocrine therapy.

Kate *et al.* (39) trained breast cancer survivability prediction models based on the cancer staging system from the SEER dataset. Previous studies found that the 5-year survival of patients with *in situ* breast cancer was as high as 99.42%, while that of patients with distant metastasis (DM) was only 36.71%. Compared with the traditional combined models trained on all summary stages, stage-specific predictive models trained separately on each summary stage are the most suitable models to predict survivability for a specific stage, although both showed poor predictive performance for survivability in patients with distant metastases.

Gene-expression assays

21-gene recurrence score (RS) assay

To quantify the likelihood of distant recurrence (DR) in tamoxifen-treated patients with node-negative, ER-positive breast cancer, Paik *et al.* (40) developed a 21-gene RS assay based on a reverse-transcriptase (RT)-polymerase-chain-reaction (PCR). Sixteen genes related to proliferation, invasion, survival, ER, and HER2, together with five reference genes were used in a prospectively defined algorithm to calculate the RS and determine a risk group (low, intermediate, or high) for each patient. In the NSABP B-14 trial (40), the DR rate in the low-risk group was significantly lower than that in the high-risk group (P<0.001), and RS could be used as a continuous predictor to provide an accurate estimate of the risk of DR in individual patients. In the NSABP B-20 trial, patients with high-RS (≥31) tumors who received cyclophosphamide, methotrexate and fluorouracil (CMF) or methotrexate and fluorouracil (MF) chemotherapy combined with tamoxifen had a 27.6% decrease in 10-year DR rate compared with those treated with tamoxifen alone. In contrast, patients with low-RS (<18) tumors derived minimal benefit from chemotherapy treatment (41). A prospective study of TAILORx (42) confirmed that endocrine therapy was non-inferior to chemotherapy plus tamoxifen treatment for patients with intermediate-RS [11–25] tumors in the analysis of invasive DFS. Even in the subgroup at high-risk clinical features, there was no evidence suggesting any chemotherapy benefit in the patients with intermediate-RS. Exploratory analysis indicated that chemotherapy was associated with a lower rate of DR than endocrine therapy for the women 50 years of age or younger, if the RS was 16–25. However, it was not clear whether this benefit was due to the cytotoxic effects of chemotherapy or the antiestrogenic effect associated with premature menopause induced by chemotherapy.

Similar to the association between RS and risk for DR, there also existed a significant association between RS and the risk of local regional recurrence (LRR) in ER-positive, node-negative patients from NSABP B-14 and B-20 clinical trials (43). Interestingly, the association in placebo-treated patients and CMF chemotherapy plus tamoxifen-treated patients was less robust than that in tamoxifen-treated patients, reflecting the differential effect of chemotherapy in reducing LRR by RS classification, similar to that shown with DR (41). In addition, unlike the straightforward association between RS and LRR in patients who underwent a mastectomy, the association

pattern between RS and LRR in patients who received radiotherapy after BCS was affected by age in the low RS group, suggesting that the effect of radiotherapy may not be uniform across RS categories but that radiotherapy may be more effective as RS increase. Several ongoing clinical trials, such as the Lumina trial, the Precision trial, and the IDEA study, are attempting to validate the clinical utility of incorporating genomic-based prognostic information into regional treatment decisions for low-risk, mostly node-negative breast cancer patients (44).

A retrospective analysis (45) of the SWOG8814 trial demonstrated that RS was also an independent predictor in ER-positive, node-positive breast cancer. The RS was highly prognostic for DFS within the tamoxifen-alone group. However, for those surviving beyond 5 years, the RS was no longer prognostic (HR =0.86; P=0.80), but the prognostic effect of the first 5 years persisted over the entire period. In addition, there was no benefit from cyclophosphamide, doxorubicin, and fluorouracil (CAF) chemotherapy in the low-RS tumors, but major 10-year DFS improvement for the high-RS group, adjusting for the number of positive nodes. The result of TransATAC research (46) further confirmed that RS is significantly associated with the risk of DR in postmenopausal hormone-receptor-positive patients treated with tamoxifen or anastrozole, and that high-RS tumors are more likely to benefit from chemotherapy regardless of lymph node status.

In the NSABP B-28 trial (47), there was also a significant correlation between RS and LRR rates in ER-positive and node-positive patients who received both doxorubicin/cyclophosphamide (AC) or AC plus paclitaxel (ACP) chemotherapy and tamoxifen treatment. Subgroup analysis showed that for patients with 4 or more positive nodes, RS was statistically significantly associated with risk of LRR, while there was a statistically nonsignificant trend in patients with 1 to 3 positive nodes. However, in another SWOG8814 study (44), there was no significant difference by RS in the 10-year rates of LRR among the patients with 4 or more positive nodes who received the mastectomy without radiotherapy (25.9% vs. 27.0%; P=0.27). Therefore, these findings suggest that RS can be used with accepted clinical variables to assess the risk of LRR during radiotherapy decision-making and identify patients who are more suitable for general radiotherapy. The ongoing TailorRT study will explore the safety of receiving radiotherapy in patients with 1 to 3 positive nodes in the low-RS subgroup. Until the results of TailorRT, radiotherapy decision-making will continue to be controversial for patients with low lymph node tumor load and low invasive tumor characteristics.

Recently, several studies have applied RS in the decision-making of neoadjuvant chemotherapy (NACT) in patients with ER-positive and HER2-negative breast cancer (48-51). These studies showed that pathologic complete response (pCR) or clinical complete response (cCR) to NACT rarely occurs in patients with low-RS, while approximately 10% of patients with high-RS (RS >25 or RS >30) achieved pCR. NACT decision-making based on RS can effectively avoid overtreatment of advanced patients with low-RS. However, in the study of TransNEOS (52), low-RS (RS <18) tumors achieved a higher proportion of clinical response (P<0.001 vs. RS ≥31) for ER-positive and node-negative postmenopausal patients. Bear et al. (53) selected neoadjuvant therapy for ER-positive patients according to RS grouping. Among them, patients with RS <11 received hormone therapy, RS ≥26 group received chemotherapy, and those with intermediate RS [11–25] were randomized to be treated with hormone therapy or chemotherapy. The results showed that the clinical response rate was significantly correlated with the RS grouping in each group (P=0.049), indicating that RS could be used to guide the neoadjuvant treatment decision for ER-positive and locally advanced postmenopausal breast cancer.

In conclusion, 21-GA provides early-stage prognostic information and the likelihood of benefit from adjuvant therapy for hormone-receptor-positive breast cancers, while no recognition of advanced DM has been observed (54). Although its use has been standardized and incorporated into a variety of international clinical guidelines, the high cost of the test still limited the prevalence of 21-GA. The pros and cons of 21-GA and other gene expression assays are presented in the *Table 1*.

70-gene prognostic signature

van't Veer et al. (55) used DNA microarray analysis and hierarchical clustering algorithm to identify a gene expression signature strongly predictive of a short interval to distant metastases in node-negative patients. The biosignature called 70-GS, was composed of 70 genes related to cell cycle, invasion, metastasis, and angiogenesis, distinguished patients with "good prognosis" and those with "poor prognosis". Besides, 70-GS effectively identified high-risk populations that might benefit from adjuvant therapy and outperformed clinical variables such as tumor size, histological grade, vascular invasion, age, and ER status in predicting the likelihood of DM within 5 years. Further research confirmed

Table 1 The applicable crowd, advantages and disadvantages and NCCN Guidelines Recommendation of gene expression assays

Model	Applicable crowd	Advantages	Disadvantages	NCCN guideline recommendations
21-GA	ER (+), node (+/−)	(I) Predict the 10-year risk of DR and LRR; (II) Predict the benefit of chemotherapy and guide patients' chemotherapy decision	(I) No recognition of the risk of late-stage events; (II) Lack of guiding significance for the chemotherapy decision in patients with intermediate-RS; (III) Lack of effective evidence to guide regional radiotherapy decision-making; (IV) High cost	(I) Class 1 evidence for patients with node (−) and postmenopausal patients with node (1–3+); (II) Class 2A evidence for premenopausal patients with node (1–3+)
70-GS	ER (+/−), node (+/−)	(I) Predict the 5-year risk of DM and LRR; (II) Guide chemotherapy decisions in patients with intermediate-RS; (III) Instruct chemotherapy decisions in patients with high-risk clinical features and low-risk gene-expression profiles	(I) Fail to effectively predict late tumor events; (II) The cost-effectiveness was observed in patients with high-risk clinical features, but fail in those with negative nodes; (III) Fresh frozen tissue samples require higher quality; (IV) High cost	(I) Class 1 evidence for patients with node (−) and patients with node (1–3+)
PAM50	All patients	(I) Assign an intrinsic subtype to patients; (II) Predict the 10-year risk of DR and LR; (III) Instruct the prolonged endocrine therapy for 5–10 years after diagnosis; (IV) Better distinguish intermediate- and high-risk groups than 21-GA; (V) Can be performed in any qualified pathology laboratory	(I) Lack of effective evidence to guide chemotherapy decision; (II) Cannot predict the benefit of radiotherapy, and need further verification; (III) High cost	(I) Class 2A evidence for patients with node (−) and patients with node (1–3+)
Endopredict	ER (+), HER2(−)	(I) Combine genomics and clinicopathological information; (II) Predict the early- and late-stage risk of DR and LR; (III) Instruct the decision-making of prolonged endocrine therapy and combining chemotherapy; (IV) Portable prognostic platform and sample type	(I) Lack of effective evidence on predicting the benefit of chemotherapy; (II) Not suitable for customized local treatment regimens	(I) Class 2A evidence for patients with node (−) and patients with node (1–3+)
BCI	ER (+), node (+/−)	(I) Unique advantages in late-stage prognostic prediction; (II) Important guiding significance for prolonged adjuvant endocrine therapy	(I) Lack of effective evidence on predicting the benefit of chemotherapy	(I) Class 2A evidence for patients with node (−) and patients with node (1–3+)

ER (+), estrogen receptor positive; ER (+/−), estrogen receptor positive or negative; HER2 (−), human epidermal growth factor receptor 2; node (+/−), lymph node positive or negative; node (1–3+), 1 to 3 positive lymph nodes; 21-GA, 21-gene assay; 70-GS, 70-gene signature; DR, distant recurrence; LRR, local regional recurrence; LR, local recurrence; BCI, breast cancer index; NCCN, National Comprehensive Cancer Network.

that 70-GS was also a strong predictor of DM in patients with positive nodes (HR =4.5, P<0.001) (56). In addition, 70-GS was significantly correlated with the risk of locoregional recurrence (LRR) (P<0.001) as well. However, the effect of RS on LRR did depend on the time since diagnosis, with a strong effect in the first 5 years (HR =2.59), and no significant relation at longer follow-up. What's more, several retrospective analyses (57-59) and a community-based feasibility study (Raster) (60) have verified the prediction performance of 70-GS. Similar to 21-GA, 70-GS predicted recurrence risk only within 5 years after diagnosis and failed to identify late-stage events effectively (54).

The MINDACT Trial provided prospective, randomized evidence of clinical efficacy of 70-GS, identifying a significant proportion of patients with discordant risk results between clinical features and gene-expression profiles in early-stage breast cancer (3). Cardoso et al. (61) confirmed that for patients with high-risk clinical features and low-risk gene-expression profiles, there was no significant difference in 5-year distant metastasis-free survival (DMFS) between patients treated with or without chemotherapy (P=0.27). In this subgroup, patients who did not receive chemotherapy could also got an extremely high rate of 5-year DMFS (94.4%), which was only 1.5% lower than those treated with chemotherapy. Approximately 46% of patients with high-risk clinical characterizes tumor may not need chemotherapy. Notably, even for patients with positive nodes [1–3], there was no statistical difference in DMFS between the two treatment groups. However, there was also no meaningful difference in the 5-year DMFS between patients with high-risk clinical features and low-risk gene-expression profiles when they received chemotherapy or not (P=0.66). Therefore, for patients with high clinical risk, the addition of 70-GS information based on traditional clinical and pathological factors can provide valuable information for the selection of patients who may benefit from adjuvant chemotherapy, while the results of 70-GS do not provide evidence for chemotherapy recommendations chemotherapy for patients with low clinical risk. In the exploratory analysis of the potential age effect, the benefit of chemotherapy on high clinical risk and low genetic risk was only observed in women under 50 years of age, which may be related to chemotherapy-induced ovarian function inhibition. Further study is required for younger (62).

In a prospective, observational, multicenter study, Kuijer et al. (63) assessed the impact of 70-GS on chemotherapy decisions in patients with early-stage, ER-positive breast cancer. After disclosure of the 70-GS result, the preliminary advice was changed in 51% of patients who received a recommendation before testing, and the final chemotherapy recommendation of the physician was consistent with the 70-GS result in 96% of patients. In addition, the PROMIS trial (64) evaluated the effect of 70-GS on the chemotherapy decision in a patient population with an intermediate-RS [18–30] derived from 21-GA. It showed that 70-GS classification was significantly associated with chemotherapy decisions of clinicians (P<0.001). The adjuvant treatment recommendations for 36% of patients were changed after receiving the results of 70-GS, and physicians reported having greater confidence in their treatment recommendations. Therefore, the 70-GS provided clinicians with clinically actionable information and increased their confidence in degraded care.

It's worth noting that the cost-effectiveness of 70-GS was observed to be superior to clinical evaluation only in patients with ER-positive, HER2-negative, and high-risk clinical features (65), while no significant advantage was seen in patients with negative nodes (66).

PAM50 risk of recurrence (ROR) score

Parker et al. (67) used microarray and quantitative RT-PCR to identify significant clusters representing different "intrinsic" subtypes and develop a 50-gene subtype predictor (PAM50), and then calculated a ROR score. For patients with node-negative tumors, the intrinsic subtype model (ROR-S) showed significantly better prediction performance than the clinical variable model but worse than the combined model (ROR-C), including subtype and clinical variables. ROR-C score was linearly associated with the 5-year recurrence rate. In addition, the ROR-S predicted NACT efficacy effectively with a negative predictive value for pCR of 97%. The pCR rate was positively correlated with ROR-S score, but there is a platform in the high-ROR-S groups, suggesting chemotherapy resistance in high-risk tumors. Nielsen et al. (68) demonstrated that PAM50 reassigned some clinical ER-positive patients to non-luminal subtypes and provided more prognostic information for long-term survival in tamoxifen-treated ER-positive patients than standard clinical variables.

In the TransATAC trial (69), the ROR score added significant prognostic information for 10-year DR beyond that of RS and clinical treatment score (CTS), and showed better discrimination performance between moderate-risk

and high-risk groups than RS. Similarly, in the ABCSG-8 trial (70), for patients with early-stage, ER-positive tumors and received 5 years endocrine therapy, the continuous ROR score and the ROR-based risk groups significantly added prognostic information for distant recurrence-free survival (DRFS) to the clinical predictors (P<0.0001). What's more, PAM50 assigned an intrinsic subtype to all patients based on the nearest centroid. DRFS was significantly higher in LumA patients compared with LumB (P<0.0001), and LumA/LumB subtypes add a significant amount of additional prognostic information to the clinical predictors (P<0.0001), regardless of lymph node status. The ABCSG-8 trial together with the ATAC trial constituted the first-level evidence of the clinical effectiveness of the PAM50 in predicting the risk of DR in postmenopausal women with ER-positive early-stage tumors (70). The combined analysis of these two trials showed that the ROR score also had clinically meaningful prognostic significance in the late follow-up period (5–10 years), and helped select the risk population who potentially benefit from prolonged hormone therapy beyond 5 years (71).

In the DBCG-77B trial (72), a highly significant association was observed between intrinsic subtypes and cyclophosphamide (C) or CMF treatment for DFS in premenopausal high-risk patients. The risk of DFS was significantly reduced in basal-like and LumB subtypes, while no benefit was seen in HER2-enriched or LumA subtypes. In addition, DBCG-89D trial (73) showed that patients with HER2-enriched or basal-like tumors benefit from adding anthracyclines (CEF) to CMF chemotherapy, which was associated with the increase of ROR score. In patients with LumA or LumB subtypes, the persistent ROR score did not show any association with anthracycline chemotherapy benefit. In addition, in the MONALEESA trial (74), all subtypes except basal-like demonstrated significant progression-free survival (PFS) benefit with ribociclib.

ROR score also provided prognostic information for local recurrence (LR). In the ABCSG-8 trial (75), the PAM50 ROR score and intrinsic subtypes effectively identified a genomic low-risk population within the group of patients with clinically low-risk of LR (super low risk of LR). However, no prognostic difference was observed between patients with and without radiotherapy, suggesting that PAM50 could not predict the benefit of radiotherapy. Due to the limited number of patients and LR events, the effect on LR of PAM50 still needs to be further verified in larger cohorts with more events.

The PAM50 ROR score and subtype classifier predicted the survival and chemotherapy benefits of breast cancer patients better than 21-GA RS and clinical variables. Despite its high cost, PAM50 can be performed in any qualified pathology laboratory, laying the foundation for its widespread application.

EndoPredict (EP) score

Filipits *et al.* (76) developed and defined a risk score EP for patients with ER-positive, HER2-negative, early-stage tumors, consisting of eight tumor-related and three reference genes. Combined with nodal status and tumor size, EP was transformed into a comprehensive risk score, EPclin. In the ABCSG-6/8 trials, with the differential contribution of proliferation genes and ER signaling genes, continuous EP provides additional prognostic information for identifying early (5 years) and late (>5 years) DM, independent of Adjuvant! Online, ER, and Ki67 (76,77). Besides, the EPclin score outperformed all conventional clinicopathologic risk factors in predicting late-stage recurrence events, identifying high-risk patients who potentially benefit from extended hormone therapy or other adjuvant therapies (76). Further study of ABCSG-6/8 trials verified the value of the EPclin score for late DR (10- or 15-year), regardless of lymph node status (78).

The prognostic significance of EPclin in DR for premenopausal and postmenopausal patients with ER-positive, HER2-negative tumors has been demonstrated (79,80). On this basis, Sestak *et al.* (81) found that patients with high EPclin score who received chemotherapy had a significantly lower 10-year DR rate than those received endocrine therapy alone, while no difference in the risk of DR was observed between the two treatment groups with a low EPclin score (<3.3). The positive interaction between the EPclin score and the treatment groups further emphasized the potential benefit of chemotherapy in patients with high EPclin score, independent of lymph node status. In addition, Bertucci *et al.* (82) confirmed that the EP classification was associated with pCR after NACT with anthracycline (P<0.001), with the high EP group more likely to obtain pCR. In the ABCSG-34 trial (83), EP score significantly predicted the benefits of NACT and neoadjuvant endocrine therapy (NET) in patients with ER-positive, HER2-negative, early-stage tumors. Patients with low EP score treated with NACT were unlikely to benefit from NACT, while a high EP score predicted resistance to NET.

The EP score was also an independent predictor of local

relapse-free survival (LRFS). In the ABCSG-8 trial (84), the EP score identified genomic low-risk individuals in clinically low- to intermediate-risk groups for LR. However, the effect on LR of EP was only observed within the group of patients receiving RT, while in women who did not receive RT after BCT, the LR rate was similar between the EP high- as well as EP low-risk group. Moreover, in patients at high risk for EP, radiotherapy did not result in fewer LR events than mastectomy and BCS, suggesting that EP may not be suitable for customized local treatment regimens.

The EP and EPclin scores provided prognostic information for early- and late- stage events in patients with ER-positive and HER2-negative tumors. The EPclin score combined with clinicopathological parameters was associated with better prognosis prediction and risk stratification. The portability of the prognostic platform and sample type is an essential feature of the clinical application of the EP.

Breast cancer index (BCI)

Jerevall et al. (85) developed a continuous risk model based on molecular grade index (MGI) and HOXB13: IL17Br (H:I) ratio to estimate recurrence risk at the individual level for patients with ER-positive, node-negative tumors. The MGI is a five-gene predictor that recapitulates tumor grade and/or proliferation and is highly prognostic in patients with ER-positive tumors. H:I is prognostic for early and late DR and is predictive of adjuvant and extended adjuvant hormonal benefit in patients with early-stage hormone-receptor-positive, node-negative tumors, independent of tumor grade and/or proliferation. The dichotomous index combined MGI and H: I outperformed either index alone in predicting the ROR (86). Similarly, continuous BCI also had significant prognostic efficacy in the untreated population, with a larger proportion of patients classified into the low-risk group for recurrence, suggesting that additional chemotherapy was unnecessary (85). NCIC CTG MA.14 Study (87) further confirmed that BCI has a strong prognostic effect on RFS in postmenopausal patients with early-stage tumors treated with tamoxifen alone or in combination with octreotide. Even in patients with positive lymph nodes, half were classified as low risk by the BCI.

Stockholm randomized controlled cohorts (88) showed that continuous BCI was the most significant predictor to assess early- and late-DR risk in patients with ER-positive, node-negative tumors, and treated with hormone therapy (P=0.0005), which assisted chemotherapy decisions at diagnosis and extended adjuvant endocrine therapy

decisions beyond 5 years. Similarly, in the Trans-aTTom trial (89), the patients with high BCI score, who have a hormone-receptor-positive, node-positive tumor, also benefited significantly from 10- and 5-year tamoxifen treatment, with an absolute risk reduction of 10.2% based on relapse-free interval (RFI) (P=0.027). The interaction between prolonged tamoxifen treatment and BCI was statistically significant (P=0.012), further confirming the potential benefit of prolonged tamoxifen treatment.

Other special gene-expression assays

Wang et al. (90) used the Affymetrix U133A gene chips to identify a 76-gene signature, consisting of 60 ER-positive and 16 ER-negative genes for patients with node-negative breast cancer who did not receive systemic adjuvant therapy. In the analysis of the function of 76 genes, although genes related to cell death, proliferation and transcriptional regulation were found in both groups of patients stratified by ER status, there was no overlap between the genes of the two group, indicating that the extent of heterogeneity and the underlying mechanisms for disease progression may be different for the two ER-based subgroups of breast cancer patients.

To solve the multicollinearity of genes and improve the generalization ability of the model, Hikichi et al. (91) applied the same dataset as which Wang et al. (90) used to build 76-Gene signature to develop a simple but robust feature selection method using a correlation-centered approach. They obtained a 12-gene set with both high predictive and generalization abilities and constructed a prediction model for 5-year DM in patients with early-stage breast cancer, of which the prediction efficiency was similar to that of Rotterdam 76-gene assay.

For TNBC patients who received adjuvant chemotherapy after curative surgery, Park et al. (92) found that there were 11, 7, 9 genes included in the final prediction models for RFS, DRFS, and OS, respectively, using NanoString expression assay and systematic gradient LASSO algorithm. These models divided patients into high-risk and low-risk groups, and provided significant prognostic ability together with the TNM staging system.

Most of the prognosis-predictive molecular scores for breast cancer are applicable to only limited disease subtypes. To address this issue, Shimizu et al. (93) obtained a minimal genome consisting of 23 genes by analyzing the entire group of human protein-coding genes. Molecular prognostic score (mPS) was calculated by a linear weighting of the 23 genes. mPS was not only suitable for both RNA-

2826

Min et al. Advancement of prognostic models in breast cancer

sequencing and microarray datasets, but also capable of stratifying OS and DFS simultaneously. In addition, mPS can further enlist patients into different risk groups for precise stratification for different PAM50 intrinsic subtypes, age, pathological subtypes, and TNM stages.

ML model (image)

Radiomics signature
Park *et al.* (94) developed a radiomics signature based on preoperative magnetic resonance imaging (MRI) to estimate DFS in patients with invasive breast cancer. The radiomics nomogram incorporating the radiomics signature and clinicopathological findings showed better predictive performance for DFS than the clinicopathological or Rad-score-only nomograms. Similarly, Yu *et al.* (95) developed a dynamic contrast-enhanced (DCE)-MRI radiomic signature for preoperative identification of axillary lymph node metastasis (ALNM) and assessed individual DFS in patients with early-stage breast cancer. It was found that for both ALNM and DFS, clinical-radiomic nomogram combining the clinical characteristics and radiomic signatures showed better clinical predictive power as well. Moreover, Koh *et al.* (96) confirmed that the Rad score derived from the three-dimensional radiomic features of breast MRI was also a prognostic factor for predicting systemic recurrence in patients with TNBC.

Tadayyon *et al.* (97) developed a quantitative ultrasound (QUS)-based multiparametric classifier incorporating texture and image quality features that account for tumor core and a 5 mm thick surrounding margin. The classifier was demonstrated to be a sensitive (90%) and specific (79%) pre-treatment predictor of tumor response to NACT and 5-year RFS ($P<0.05$) in patients with locally advanced breast cancer (LABC). Furthermore, Dasgupta *et al.* (98) performed a second-pass texture analysis on QUS parameters to create high-order texture derivatives to predict the NAC response of LABC, and obtained better results than using texture features alone. In addition, Xiong *et al.* (99) extracted a multi-feature-based radiomics signature from preoperative ultrasound (US) images and calculated a radiomics score (Rad-score) to predict the DFS of invasive breast cancer. The Rad-score was significantly correlated with DFS and provided additional predictive performance independent of clinical pathology nomograph. Results of subgroup analysis showed that the Rad-score successfully discriminate prognoses in LumB ($P=0.00006$) and TNBC ($P=0.00003$), but failed in either LumA ($P=0.563$) or HER2-enriched ($P=0.109$), indicating that the ability of radiomics signature to assess DFS for invasive breast cancer vary by molecular subtype.

Pathological recurrence classifier
Klimov *et al.* (100) developed a novel ML pipeline to predict the risk of ipsilateral recurrence. For DCIS patients, at first, the patient's whole surgical H&E slides were transformed into prognostically information tissue classes, through the random forest annotation classifier manually trained by pathologists. Then, tissue architecture features and features of the spatial relationship between these tissue classes are extracted and compiled into a "full sliding" feature set. Finally, a random forest recurrence risk classifier is trained by combining the patient's digitized whole slide image (WSI) and clinical pathological long-term result data. As a result, the recurrence classifier significantly predicted the 10-year recurrence risk of DCIS and provided predictive value for the long-term outcome of radiotherapy after BCS in patients with different risk groups.

Discussion

Breast cancer is a group of heterogeneous diseases with a variety of molecular subtypes. Conventional clinicopathological factors and their combined models are often difficult to accurately predict the prognosis and treatment benefits of individual breast cancer patients. To better define risk stratification, whole-genome studies have created multiple prognostic genetic signatures for breast cancer. The first generation of polygenic testing tools, represented by 21-GA and 70-GS, provide early survival prognosis and treatment benefit prediction independent of clinicopathological factors for breast cancer patients. The second generation of multigene testing tools, represented by PAM50, EPclin, and BCI, combined genomic information and clinical variables to build a composite model. Their risk stratification and prediction performance were better than the first generation, and they could achieve the prediction of long-term survival events. AJCC, National Comprehensive Cancer Network (NCCN), and other guidelines have recommended that the above genetic testing tools could be applied in clinical practice. However, the application of polygenic testing is still limited due to high cost, insufficient technology, poor reproducibility and other reasons. Artificial intelligence technology facilitated the emerge of prognostic models relying on ML. High-latitude data of imaging and pathological images have provided more information for prognosis prediction of breast cancer.

But such models are still in the development stage. At present, there is still a lack of economic and practical prognostic tools for clinical application, especially for ER-negative breast cancer patients. Further studies need to be conducted for the validation of existing models and the development of new efficient prognostic models including both clinical and genomic data. In this article, we merely searched the PubMed database and summarized the classic prognostic models of breast cancer. Researches on prognostic and verification models that combine multidisciplinary fields with new technologies like artificial intelligence need to be conducted.

In conclusion, patients with breast cancer should be stratified based on genomics information and clinical characteristics simultaneously to identify low-risk patients who may avoid unnecessary systemic treatment, especially for patients with high-risk clinical features and low-risk gene-expression profiles. The detailed interpretation of imaging or pathological image will provide more accurate information for the patient's prognosis prediction.

Acknowledgments

Footnote

References

1. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100:8418-23.

2. Giuliano AE, Edge SB, Hortobagyi GN. Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. Ann Surg Oncol 2018;25:1783-5.

3. Audeh W, Blumencranz L, Kling H, et al. Prospective Validation of a Genomic Assay in Breast Cancer: The 70-gene MammaPrint Assay and the MINDACT Trial. Acta Med Acad 2019;48:18-34.

4. Phung MT, Tin Tin S, Elwood JM. Prognostic models for breast cancer: a systematic review. BMC Cancer 2019;19:230.

5. Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. Br J Cancer 1982;45:361-6.

6. Todd JH, Dowle C, Williams MR, et al. Confirmation of a prognostic index in primary breast cancer. Br J Cancer 1987;56:489-92.

7. D'Eredita' G, Giardina C, Martellotta M, et al. Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution. Eur J Cancer 2001;37:591-6.

8. Brown J, Jones M, Benson EA. Comment on the Nottingham Prognostic Index. Breast Cancer Res Treat 1993;25:283.

9. Balslev I, Axelsson CK, Zedeler K, et al. The Nottingham Prognostic Index applied to 9,149 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). Breast Cancer Res Treat 1994;32:281-90.

10. Blamey RW, Pinder SE, Ball GR, et al. Reading the prognosis of the individual with breast cancer. Eur J Cancer 2007;43:1545-7.

11. Blamey RW, Ellis IO, Pinder SE, et al. Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999. Eur J Cancer 2007;43:1548-55.

12. Al jarroudi O, Zaimi A, Brahmi SA, et al. Nottingham Prognostic Index is an Applicable Prognostic Tool in Non-Metastatic Triple-Negative Breast Cancer. Asian Pac J Cancer Prev 2019;20:59-63.

2828

Min et al. Advancement of prognostic models in breast cancer

13. Albergaria A, Ricardo S, Milanezi F, et al. Nottingham Prognostic Index in triple-negative breast cancer: a reliable prognostic tool? BMC Cancer 2011;11:299.

14. Kwon J, Eom KY, Koo TR, et al. A Prognostic Model for Patients with Triple-Negative Breast Cancer: Importance of the Modified Nottingham Prognostic Index and Age. J Breast Cancer 2017;20:65-73.

15. Van Belle V, Van Calster B, Brouckaert O, et al. Qualitative assessment of the progesterone receptor and HER2 improves the Nottingham Prognostic Index up to 5 years after breast cancer diagnosis. J Clin Oncol 2010;28:4129-34.

16. Winzer KJ, Buchholz A, Schumacher M, et al. Improving the Prognostic Ability through Better Use of Standard Clinical Data - The Nottingham Prognostic Index as an Example. PLoS One 2016;11:e0149977.

17. Rakha EA, Soria D, Green AR, et al. Nottingham Prognostic Index Plus (NPI+): a modern clinical decision making tool in breast cancer. Br J Cancer 2014;110:1688-97.

18. Green AR, Soria D, Stephen J, et al. Nottingham Prognostic Index Plus: Validation of a clinical decision making tool in breast cancer in an independent series. J Pathol Clin Res 2016;2:32-40.

19. Green AR, Soria D, Powe DG, et al. Nottingham prognostic index plus (NPI+) predicts risk of distant metastases in primary breast cancer. Breast Cancer Res Treat 2016;157:65-75.

20. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 2001;19:980-91.

21. Ozanne EM, Braithwaite D, Sepucha K, et al. Sensitivity to input variability of the Adjuvant! Online breast cancer prognostic model. J Clin Oncol 2009;27:214-9.

22. de Glas NA, van de Water W, Engelhardt EG, et al. Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study. Lancet Oncol 2014;15:722-9.

23. Olivotto IA, Bajdik CD, Ravdin PM, et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. J Clin Oncol 2005;23:2716-25.

24. Mook S, Schmidt MK, Rutgers EJ, et al. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online Adjuvant! program: a hospital-based retrospective cohort study. Lancet Oncol 2009;10:1070-6.

25. Campbell HE, Taylor MA, Harris AL, et al. An investigation into the performance of the Adjuvant!

26. Hajage D, de Rycke Y, Bollet M, et al. External validation of Adjuvant! Online breast cancer prognosis tool. Prioritising recommendations for improvement. PLoS One 2011;6:e27446.

27. Jung M, Choi EH, Nam CM, et al. Application of the adjuvant! Online model to Korean breast cancer patients: an assessment of prognostic accuracy and development of an alternative prognostic tool. Ann Surg Oncol 2013;20:2615-24.

28. Yao-Lung K, Dar-Ren C, Tsai-Wang C. Accuracy validation of adjuvant! online in Taiwanese breast cancer patients--a 10-year analysis. BMC Med Inform Decis Mak 2012;12:108.

29. Bhoo-Pathy N, Yip CH, Hartman M, et al. Adjuvant! Online is overoptimistic in predicting survival of Asian breast cancer patients. Eur J Cancer 2012;48:982-9.

30. Wishart GC, Azzato EM, Greenberg DC, et al. PREDICT: a new UK prognostic model that predicts survival following surgery for invasive breast cancer. Breast Cancer Res 2010;12:R1.

31. Wishart GC, Bajdik CD, Azzato EM, et al. A population-based validation of the prognostic model PREDICT for early breast cancer. Eur J Surg Oncol 2011;37:411-7.

32. Wishart GC, Bajdik CD, Dicks E, et al. PREDICT Plus: development and validation of a prognostic model for early breast cancer that includes HER2. Br J Cancer 2012;107:800-7.

33. Wishart GC, Rakha E, Green A, et al. Inclusion of KI67 significantly improves performance of the PREDICT prognostication and prediction model for early breast cancer. BMC Cancer 2014;14:908.

34. Drukker CA, Nijenhuis MV, Bueno-de-Mesquita JM, et al. Optimized outcome prediction in breast cancer by combining the 70-gene signature with clinical risk prediction algorithms. Breast Cancer Res Treat 2014;145:697-705.

35. Candido Dos Reis FJ, Wishart GC, Dicks EM, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res 2017;19:58.

36. Gray E, Marti J, Brewster DH, et al. Independent validation of the PREDICT breast cancer prognosis prediction tool in 45,789 patients using Scottish Cancer Registry data. Br J Cancer 2018;119:808-14.

37. Bremer T, Whitworth PW, Patel R, et al. A Biological

Online prognostic programme in early breast cancer for a cohort of patients in the United Kingdom. Br J Cancer 2009;101:1074-84.

Signature for Breast Ductal Carcinoma In Situ to Predict Radiotherapy Benefit and Assess Recurrence Risk. Clin Cancer Res 2018;24:5895-901.

38. Qiu Y, Wang L, Zhong X, et al. A multiple breast cancer stem cell model to predict recurrence of T1-3, N0 breast cancer. BMC Cancer 2019;19:729.

39. Kate RJ, Nadig R. Stage-specific predictive models for breast cancer survivability. Int J Med Inform 2017;97:304-11.

40. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004;351:2817-26.

41. Paik S, Tang G, Shak S, et al. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006;24:3726-34.

42. Sparano JA, Gray RJ, Makower DF, et al. Adjuvant Chemotherapy Guided by a 21-Gene Expression Assay in Breast Cancer. N Engl J Med 2018;379:111-21.

43. Mamounas EP, Tang G, Fisher B, et al. Association between the 21-gene recurrence score assay and risk of locoregional recurrence in node-negative, estrogen receptor-positive breast cancer: results from NSABP B-14 and NSABP B-20. J Clin Oncol 2010;28:1677-83.

44. Woodward WA, Barlow WE, Jagsi R, et al. Association Between 21-Gene Assay Recurrence Score and Locoregional Recurrence Rates in Patients With Node-Positive Breast Cancer. JAMA Oncol 2020;6:505-11.

45. Albain KS, Barlow WE, Shak S, et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 2010;11:55-65.

46. Dowsett M, Cuzick J, Wale C, et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol 2010;28:1829-34.

47. Mamounas EP, Liu Q, Paik S, et al. 21-Gene Recurrence Score and Locoregional Recurrence in Node-Positive/ER-Positive Breast Cancer Treated With Chemo-Endocrine Therapy. J Natl Cancer Inst 2017;109:djw259.

48. Morales Murillo S, Gasol Cudos A, Veas Rodriguez J, et al. Selection of neoadjuvant treatment based on the 21-GENE test results in luminal breast cancer. Breast 2021;56:35-41.

49. Pease AM, Riba LA, Gruner RA, et al. Oncotype

DX® Recurrence Score as a Predictor of Response to Neoadjuvant Chemotherapy. Ann Surg Oncol 2019;26:366-71.

50. Kantor O, Barrera E, Kopkash K, et al. Are we Overtreating Hormone Receptor Positive Breast Cancer with Neoadjuvant Chemotherapy? Role of OncotypeDx® for Hormone Receptor Positive Patients Undergoing Neoadjuvant Chemotherapy. Ann Surg Oncol 2019;26:3232-9.

51. Pivot X, Mansi L, Chaigneau L, et al. In the era of genomics, should tumor size be reconsidered as a criterion for neoadjuvant chemotherapy? Oncologist 2015;20:344-50.

52. Iwata H, Masuda N, Yamamoto Y, et al. Validation of the 21-gene test as a predictor of clinical response to neoadjuvant hormonal therapy for ER+, HER2-negative breast cancer: the TransNEOS study. Breast Cancer Res Treat 2019;173:123-33.

53. Bear HD, Wan W, Robidoux A, et al. Using the 21-gene assay from core needle biopsies to choose neoadjuvant therapy for breast cancer: A multicenter trial. J Surg Oncol 2017;115:917-23.

54. Sgroi DC, Sestak I, Cuzick J, et al. Prediction of late distant recurrence in patients with oestrogen-receptor-positive breast cancer: a prospective comparison of the breast-cancer index (BCI) assay, 21-gene recurrence score, and IHC4 in the TransATAC study population. Lancet Oncol 2013;14:1067-76.

55. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530-6.

56. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999-2009.

57. Mook S, Schmidt MK, Viale G, et al. The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study. Breast Cancer Res Treat 2009;116:295-302.

58. Buyse M, Loi S, van't Veer L, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006;98:1183-92.

59. Bueno-de-Mesquita JM, Linn SC, Keijzer R, et al. Validation of 70-gene prognosis signature in node-negative breast cancer. Breast Cancer Res Treat 2009;117:483-95.

60. Bueno-de-Mesquita JM, van Harten WH, Retel VP, et al. Use of 70-gene signature to predict prognosis of

2830

Min et al. Advancement of prognostic models in breast cancer

patients with node-negative breast cancer: a prospective community-based feasibility study (RASTER). Lancet Oncol 2007;8:1079-87.

61. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. N Engl J Med 2016;375:717-29.

62. Piccart M, van 't Veer LJ, Poncet C, et al. 70-gene signature as an aid for treatment decisions in early breast cancer: updated results of the phase 3 randomised MINDACT trial with an exploratory analysis by age. Lancet Oncol 2021;22:476-88.

63. Kuijer A, Straver M, den Dekker B, et al. Impact of 70-Gene Signature Use on Adjuvant Chemotherapy Decisions in Patients With Estrogen Receptor-Positive Early Breast Cancer: Results of a Prospective Cohort Study. J Clin Oncol 2017;35:2814-9.

64. Tsai M, Lo S, Audeh W, et al. Association of 70-Gene Signature Assay Findings With Physicians' Treatment Guidance for Patients With Early Breast Cancer Classified as Intermediate Risk by the 21-Gene Assay. JAMA Oncol 2018;4:e173470.

65. Retèl VP, Byng D, Linn SC, et al. Cost-effectiveness analysis of the 70-gene signature compared with clinical assessment in breast cancer based on a randomised controlled trial. Eur J Cancer 2020;137:193-203.

66. Bonastre J, Marguet S, Lueza B, et al. Cost effectiveness of molecular profiling for adjuvant decision making in patients with node-negative breast cancer. J Clin Oncol 2014;32:3513-9.

67. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009;27:1160-7.

68. Nielsen TO, Parker JS, Leung S, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. Clin Cancer Res 2010;16:5222-32.

69. Dowsett M, Sestak I, Lopez-Knowles E, et al. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol 2013;31:2783-90.

70. Gnant M, Filipits M, Greil R, 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;25:339-45.

71. Sestak I, Cuzick J, Dowsett M, et al. Prediction of late

distant recurrence after 5 years of endocrine treatment: a combined analysis of patients from the Austrian breast and colorectal cancer study group 8 and arimidex, tamoxifen alone or in combination randomized trials using the PAM50 risk of recurrence score. J Clin Oncol 2015;33:916-22.

72. Jensen MB, Lænkholm AV, Nielsen TO, et al. The Prosigna gene expression assay and responsiveness to adjuvant cyclophosphamide-based chemotherapy in premenopausal high-risk patients with breast cancer. Breast Cancer Res 2018;20:79.

73. Jensen MB, Lænkholm AV, Balslev E, et al. The Prosigna 50-gene profile and responsiveness to adjuvant anthracycline-based chemotherapy in high-risk breast cancer patients. NPJ Breast Cancer 2020;6:7.

74. 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;39:1458-67.

75. Fitzal F, Filipits M, Fesl C, et al. PAM-50 predicts local recurrence after breast cancer surgery in postmenopausal patients with ER+/HER2- disease: results from 1204 patients in the randomized ABCSG-8 trial. Br J Surg 2021;108:308-14.

76. Filipits M, Rudas M, Jakesz R, et al. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. Clin Cancer Res 2011;17:6012-20.

77. Dubsky P, Brase JC, Jakesz R, et al. The EndoPredict score provides prognostic information on late distant metastases in ER+/HER2- breast cancer patients. Br J Cancer 2013;109:2959-64.

78. Filipits M, Dubsky P, Rudas M, et al. Prediction of Distant Recurrence Using EndoPredict Among Women with ER+, HER2- Node-Positive and Node-Negative Breast Cancer Treated with Endocrine Therapy Only. Clin Cancer Res 2019;25:3865-72.

79. Martin M, Brase JC, Calvo L, et al. Clinical validation of the EndoPredict test in node-positive, chemotherapy-treated ER+/HER2- breast cancer patients: results from the GEICAM 9906 trial. Breast Cancer Res 2014;16:R38.

80. Dubsky P, Filipits M, Jakesz R, et al. EndoPredict improves the prognostic classification derived from common clinical guidelines in ER-positive, HER2-negative early breast cancer. Ann Oncol 2013;24:640-7.

81. Sestak I, Martín M, Dubsky P, et al. Prediction of chemotherapy benefit by EndoPredict in patients with

breast cancer who received adjuvant endocrine therapy plus chemotherapy or endocrine therapy alone. Breast Cancer Res Treat 2019;176:377-86.

82. Bertucci F, Finetti P, Viens P, et al. EndoPredict predicts for the response to neoadjuvant chemotherapy in ER-positive, HER2-negative breast cancer. Cancer Lett 2014;355:70-5.

83. Dubsky PC, Singer CF, Egle D, et al. The EndoPredict score predicts response to neoadjuvant chemotherapy and neoendocrine therapy in hormone receptor-positive, human epidermal growth factor receptor 2-negative breast cancer patients from the ABCSG-34 trial. Eur J Cancer 2020;134:99-106.

84. Fitzal F, Filipits M, Rudas M, et al. The genomic expression test EndoPredict is a prognostic tool for identifying risk of local recurrence in postmenopausal endocrine receptor-positive, her2neu-negative breast cancer patients randomised within the prospective ABCSG 8 trial. Br J Cancer 2015;112:1405-10.

85. Jerevall PL, Ma XJ, Li H, et al. Prognostic utility of HOXB13:IL17BR and molecular grade index in early-stage breast cancer patients from the Stockholm trial. Br J Cancer 2011;104:1762-9.

86. Ma XJ, Salunga R, Dahiya S, et al. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. Clin Cancer Res 2008;14:2601-8.

87. Sgroi DC, Chapman JA, Badovinac-Crnjevic T, et al. Assessment of the prognostic and predictive utility of the Breast Cancer Index (BCI): an NCIC CTG MA.14 study. Breast Cancer Res 2016;18:1.

88. Zhang Y, Schnabel CA, Schroeder BE, et al. Breast cancer index identifies early-stage estrogen receptor-positive breast cancer patients at risk for early- and late-distant recurrence. Clin Cancer Res 2013;19:4196-205.

89. Bartlett JMS, Sgroi DC, Treuner K, et al. Breast Cancer Index and prediction of benefit from extended endocrine therapy in breast cancer patients treated in the Adjuvant Tamoxifen-To Offer More? (aTTom) trial. Ann Oncol 2019;30:1776-83.

90. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005;365:671-9.

91. Hikichi S, Sugimoto M, Tomita M. Correlation-centred

variable selection of a gene expression signature to predict breast cancer metastasis. Sci Rep 2020;10:7923.

92. Park YH, Jung HH, Do IG, et al. A seven-gene signature can predict distant recurrence in patients with triple-negative breast cancers who receive adjuvant chemotherapy following surgery. Int J Cancer 2015;136:1976-84.

93. Shimizu H, Nakayama KI. A 23 gene-based molecular prognostic score precisely predicts overall survival of breast cancer patients. EBioMedicine 2019;46:150-9.

94. Park H, Lim Y, Ko ES, et al. Radiomics Signature on Magnetic Resonance Imaging: Association with Disease-Free Survival in Patients with Invasive Breast Cancer. Clin Cancer Res 2018;24:4705-14.

95. Yu Y, Tan Y, Xie C, et al. Development and Validation of a Preoperative Magnetic Resonance Imaging Radiomics-Based Signature to Predict Axillary Lymph Node Metastasis and Disease-Free Survival in Patients With Early-Stage Breast Cancer. JAMA Netw Open 2020;3:e2028086.

96. Koh J, Lee E, Han K, et al. Three-dimensional radiomics of triple-negative breast cancer: Prediction of systemic recurrence. Sci Rep 2020;10:2976.

97. Tadayyon H, Sannachi L, Gangeh MJ, et al. A priori Prediction of Neoadjuvant Chemotherapy Response and Survival in Breast Cancer Patients using Quantitative Ultrasound. Sci Rep 2017;7:45733.

98. Dasgupta A, Brade S, Sannachi L, et al. Quantitative ultrasound radiomics using texture derivatives in prediction of treatment response to neo-adjuvant chemotherapy for locally advanced breast cancer. Oncotarget 2020;11:3782-92.

99. Xiong L, Chen H, Tang X, et al. Ultrasound-Based Radiomics Analysis for Predicting Disease-Free Survival of Invasive Breast Cancer. Front Oncol 2021;11:621993.

100. Klimov S, Miligy IM, Gertych A, et al. A whole slide image-based machine learning approach to predict ductal carcinoma in situ (DCIS) recurrence risk. Breast Cancer Res 2019;21:83.