Int J Clin Exp Med 2026;19(1):20-30
www.ijcem.com /ISSN:1940-5901/IJCEM0168457

Original Article

Evaluation and exploration of the prognostic prediction efficiency of the percentile scoring system for T1-2N0M0 breast cancer patients

Zhiyong Liu, Ran Chen

Breast Treatment Center, The First Affiliated Hospital of Gannan Medical University, Ganzhou 341000, Jiangxi, China

Received August 30, 2025; Accepted January 9, 2026; Epub January 15, 2026; Published January 30, 2026

Abstract: Objective: To construct a prognostic scoring scale by integrating the data of tumor individualized immunohistochemical characteristics, so as to assist in individualized systematic treatment decision-making. Methods: A total of 1,216 female patients with T1-2N0M0 stage breast cancer were selected as the study participants. Formalin-fixed, paraffin-embedded (FFPE) tumor specimens were examined histologically and immunohistochemically in the pathological experimental center of the First Affiliated Hospital of Gannan Medical University. Antibody staining (estrogen receptor, progesterone receptor, HER2, Ki-67, CK14, FOXA1, FOXP3, PD-L1, P53, SMA, Androgen receptor, E-cadherin, CD4, CD8, CK5/6, EGFR) were used to process the archived materials. The risk factors of death in this group were analyzed, and the prognosis score scale was constructed, and the research results were compared. Results: The factors influencing mortality with clinical and statistical significance were evaluated. Logistic regression method was used to select the 10 factors that had the most significant impact on the results, and multiple scales were compiled, including regression scale 10 (based on the identified 10 most significant factors). The survival analysis of high-risk and low-risk patients using regression scale showed that there were significant differences between these groups (P<0.00001). The curative effect of the medium and high-risk patients combined with the adjuvant chemotherapy group was evaluated. The results showed that there was a significant difference in the survival rate of the medium and high-risk patients receiving adjuvant chemotherapy (P=0.0057). The regression scale of 10-year prognosis had sufficient sensitivity (58.05%), specificity (69.47%) and effectiveness (63.76%). Conclusion: The regression prognostic scale constructed in this study contains markers with high prognostic value. Multivariate analysis of the 10-year prognostic regression scale for breast cancer has improved its accuracy and reliability. This IHC-based model provides a cost-effective and biologically comprehensive alternative to existing genomic tools, particularly valuable for risk stratification in settings with limited resources.

Keywords: Breast cancer, biomarkers, prognostic scale, breast cancer prognosis

Introduction

Breast cancer is one of the most common malignant tumors in women [1]. In the past decades, thanks to the wide application of breast X-ray screening and adjuvant and neoadjuvant system therapy, the mortality of breast cancer worldwide has shown a downward trend [2]. However, the application of adjuvant chemotherapy is now more selective, guided by predictive biomarkers and multigene assays, rather than a one-size-fits-all standard for early-stage disease [3]. In depth analysis of tumor molecular genetic characteristics (such as estrogen receptor, progesterone receptor, HER2 and Ki-67 expression levels) has important practical value for carrying out individualized prognosis evaluation and selecting adjuvant therapy, but there are still some limitations in individualized prognosis evaluation [4]. Therefore, exploring new breast cancer biomarkers is gradually becoming an important research direction. A large number of studies have shown that p53 [5], CK5/6 [6], SMA [7], p63 [8], PhH3 [9], E-cadherin [10], EGFR [11], FOXA1 [12], Androgen receptor (AR) [13], Tumor Infiltrating Lymphocytes (TILs) [14] and other indicators have predictive and/or prognostic

https://doi.org/10.62347/LULX9472

significance. Although genomic profiling tools (e.g., Oncotype DX, MammaPrint) offer refined prognosis, their high cost and limited availability restrict widespread use, particularly in developing regions. This highlights the need for robust, cost-effective prognostic models based on immunohistochemistry, which can integrate both established and novel biological insights, such as tumor-infiltrating lymphocytes, to guide therapy decisions. The purpose of this study is to reduce the phenomenon of irrational drug use, improve the effectiveness of systematic treatment of breast cancer, and develop a prognosis score scale to achieve precise treatment, and provide individualized decision-making basis for breast cancer adjuvant systemic treatment.

Materials and methods

Retrospective cancer registry data analysis

This study is a retrospective analysis of breast cancer patients treated at the First Affiliated Hospital of Gannan Medical University from 2000 to 2012. Data on recurrence-free survival and overall survival were obtained through follow-up, telephone interviews, or outpatient records. The study focused on 1,216 female patients with stage T1-2N0M0 breast cancer. Key clinical and pathological characteristics are summarized in **Table 1**. The median follow-up duration was 12 years.

This study was approved by the Ethics Committee of the First Affiliated Hospital of Gannan Medical University (approval No. LLsc-2024 No. 334, approval time: January 2, 2024). The test registration number of this experiment is LLsc-2024 No. 231.

Inclusion and exclusion criteria

Inclusion criteria: Female patients aged ≥18 years diagnosed with primary unilateral invasive breast carcinoma, stage T1-2N0M0 (AJCC 7th edition), between January 2000 and December 2012 at our institution. Patients must have undergone curative surgery, have available formalin-fixed paraffin-embedded (FFPE) tumor tissue blocks, and have complete clinical, pathological, and follow-up data (minimum follow-up of 5 years unless death occurred earlier).

Exclusion criteria: Male breast cancer; history of any prior malignancy; synchronous bilateral breast cancer; received neoadjuvant systemic therapy; incomplete immunohistochemical staining results or missing key clinical/pathological variables; lost to follow-up or with incomplete survival data.

Immunohistochemical analysis of breast cancer tissue samples

Immunohistochemistry was performed on FFPE tissue sections using standardized protocols for antibodies against ER, PR, HER2, Ki-67, CK14, FOXA1, FOXP3, PD-L1, p53, SMA, AR, E-cadherin, CD4, CD8, CK5/6, and EGFR. Details of clones, dilutions, and staining conditions are provided in **Table 2**. All slides were evaluated independently by two pathologists blinded to clinical outcomes.

Development of the percentile prognostic scoring system

The prognostic scoring system was developed as follows: Predictor selection: Univariate logistic regression was performed to identify variables associated with death outcome. Variables with $P<0.10$ were entered into a stepwise multivariate logistic regression model. Ten independent prognostic factors were retained based on statistical significance ($P<0.05$) and contribution to model discrimination (increase in area under the receiver operating characteristic curve, AuROC).

Scoring method: Regression coefficients from the final multivariate model were normalized to a 0-100-point scale. The score for each factor was calculated as: (Factor Coefficient/Sum of All Coefficients) × 100, rounded to the nearest integer. The total score represented the individual's prognostic risk.

Risk stratification: Patients were categorized into three risk groups based on total score: low risk (≤40), intermediate risk (41-60), and high risk (>60).

Specific indicators used: The ten indicators included in the scoring system were: CK14 expression, FOXP3 score of 0, tumor stage T2N0M0, E-cadherin expression, p53 expression, HER2 score of 3, CD8+ T-cell count >0, EGFR expression, tumor grade G2/G3, and

Table 1. Age of patients and clinical and morphological characteristics of tumors in the study population (n=1216)

Parameter	*n (%)*	M ± SD	Me [LQ; UQ]	Min-max
Age at surgery, years	1175 (96.63)	55.66 ± 10.90	55.00 [48.00; 63.00]	24.00-84.00
Cell density, cells/mm^2	723 (59.46)	5887.61 ± 3192.83	5329.80 [4276.00; 6826.86]	940.84-48 521.89
ER, Allred score	741 (60.94)	3.80 ± 3.94	0.00 [0.00; 8.00]	0.00-8.00
PR, Allred score	741 (60.94)	3.21 ± 3.79	0.00 [0.00; 8.00]	0.00-10.00
Ki-67, %	732 (60.20)	20.02 ± 20.96	11.92 [5.35; 26.10]	0.00-96.76
Pretreatment tumor size, cm	1109 (91.20)	2.24 ± 0.87	2.00 [1.50; 2.80]	0.50-6.00
Maximum tumor size, cm	1133 (93.17)	2.33 ± 1.00	2.20 [1.50; 3.00]	0.30-15.00
Postoperative follow-up duration, years	1191 (97.94)	12.27 ± 3.60	1.00 [10.00; 16.00]	0.00-18.00

Note: ER: estrogen receptor; PR: progesterone receptor; N (%) - the absolute number and proportion of patients with corresponding parameter data in the study population.

Table 2. Immunohistochemical methods for tumor tissue samples

Antibody	Clone	Manufacturer	Incubation time	Titer	Visualization system	Treatment
CK14	LL002	Leika	30 min	1:50	EnVision Flex	TRS 9,0
FOXA1	SP133	CMQ	30 min	1:100	EnVision Flex	TRS 9,0
FOXP3	EP 340	Epitomix	o/n	1:50	EnVision Flex	TRS 9,0
PD-L1	rmAb ZR3	GeneTech	50 min	1:100	EnVision Flex	TRS 9,0
p53	Клон DO7	DAKO	30 min	1:100	EnVision Flex	TRS 6,0
Smooth muscle actin	1A4	CMQ	30 min	1:100	EnVision Flex	TRS 9,0
Androgen receptors (AR)	AR441	DAKO	30 min	1:100	EnVision Flex	TRS 9,0
E-cadherin	M	DBS	30 min	1:40	EnVision Flex	TRS 9,0
CD4	SP35	Ventana	32 min	RTU	UltraView	CCI/96°C/S
CD8	SP57	Ventana	32 min	RTU	UltraView	CCI/96°C
CK 5/6	D5/16B4	Ventana	32 min	RTU	UltraView	CCI/96°C/S64
EGFR	3C6	Ventana	36 min	RTU	UltraView	Protease 1/8 min
Estrogen receptors (ER)	SP1	Ventana	36 min	RTU	UltraView	CCI/98°C/S64
HER2	4B5	Ventana	36 min	RTU	UltraView	CCI/96°C/S36
Ki-67	30.9	Ventana	32 min	RTU	UltraView	CCI/96°C/S64
Progesterone receptors (PR)	1E2	Ventana	24 min	RTU	UltraView DAB	CCI/96°C/S64
PD-L1	22C3	DAKO	According to the stainer protocol			

Note: n: the number of patients in the corresponding category; N: Number of patients with this parameter data.

CD4+ T-cell count >0 (see **Table 6** for coefficients and scores).

Statistical analysis method

Data are presented as mean ± standard deviation or frequency (percentage). A two-tailed *p*-value <0.05 was considered statistically significant. Logistic regression and stepwise selection were used for model building. Survival analysis was performed using Kaplan-Meier method and Cox proportional hazards models. The predictive performance was assessed using sensitivity, specificity, and AuROC. Statistical analyses were conducted using Statistica 10 and SAS JMP 11.

Result

Distribution of tumor markers in breast cancer

This study analyzed the data of 1,216 patients with T1-2N0M0 breast cancer. In the study group, women aged ≥50 years old comprised the largest group (69.1%). According to TNM system assessment, T1N0M0 patients accounted for 55.2%. T2n0m0 was confirmed in 44.8% of the cohort. The tumor size of 52.0% patients was >2 cm, 46.2% was 1-2 cm, and 1.8% was <1 cm. When analyzing the histological malignancy of the tumor, the study found that the histological grading was as follows: G1 accounted for 15.6%, G2 accounted for 47.5%,

Table 3. Distribution of tumor markers in <50 and ≥50 age groups

Marker	Age group, n/N (%)		p
	<50 years	≥50 years	
CK5	21/199 (10.6)	55/483 (11.4)	0.7529
SMA	3/70 (4.3)	6/176 (3.4)	0.7411
CK14	7/129 (5.4)	25/319 (7.8)	0.3697
p53	3/69 (4.3)	7/175 (4.0)	0.9018
E-cadherin	170/220 (77.3)	417/536 (77.8)	0.8748
EGFR	27/272 (9.9)	68/606 (11.2)	0.5680
CD4	237/266 (89.1)	482/591 (81.6)	0.0055
CD8	180/257 (70.0)	375/556 (67.4)	0.4602
FOXP3	67/111 (60.4)	106/213 (49.8)	0.0696
FOXA1	192/210 (91.4)	413/453 (91.2)	0.9127
Androgen receptors (AR)	185/286 (64.7)	392/599 (65.4)	0.8250
PD-L1	11/112 (9.8)	30/216 (13.9)	0.2909

Note: n: the number of patients in the corresponding category; N: number of patients with this parameter data.

Table 4. Stepwise logistic regression results for predicting "death outcome"

No	Factor	AuROC	AuROC change	p
1	Expression of CK14	0.612	0.112	0.0159
2	FOXP3 expression score of 0	0.635	0.023	0.0048
3	Stage T2N0M0	0.639	0.004	0.0056
4	Expression of E-cadherin	0.647	0.008	0.1405
5	Expression of p53	0.652	0.005	0.0073
6	HER2 expression score of 3	0.662	0.010	0.0123
7	CD8+ T-cell count >0	0.663	0.001	0.0351
8	Expression of EGFR	0.665	0.002	0.3987
9	Tumor grade G2, G3	0.667	0.002	0.2228
10	CD4+ T-cell count >0	0.669	0.001	0.4719

G3 accounted for 36.9%. The expression of HER2 was 70.6%, 11.5%, 9.2% and 8.7% respectively. In this study, 45.7% of female estrogen receptor (ER) positive scores were in the 7-8 range. A total of 35.4% of the female progesterone receptor (PR) positive score was in the range of 7-8 points. In terms of Ki-67 expression level, 76.9% of women reached more than 5%, and about 47.7% of patients' Ki-67 level exceeded 13%. Therefore, there were significant differences in the distribution of tumor markers in the cohort (T1-2N0M0) with similar initial clinical criteria.

In this study, the distribution of other tumor markers in patients under 50 years old and patients over 50 years old were analyzed. The term "distribution" refers to the proportion of patients with positive expression for each marker, assessed via standardized IHC scoring (e.g., Allred score for ER/PR, H-score for others) by two blinded pathologists. Localization (nuclear, cytoplasmic, membranous) was defined per marker. The results showed that there were no statistically significant differences in other tumor markers except CD4 (see **Table 3** for details).

Development of prediction scoring model

In the first stage of the assessment, a single factor analysis of death risk factors was carried out to analyze the impact of a group of indicators. According to the data obtained, the most critical factors affecting the death outcome parameters from the statistical and clinical levels were screened out. Among these factors, based on the logistic regression model, 10 factors that have the most significant impact on this variable are selected. The results of stepwise logistic regression analysis are shown in **Table 4**. Positive expression for each factor was defined as follows: CK14 (>1% cytoplasmic staining), FOXP3 score of 0 (no nuclear staining in TILs), E-cadherin (loss of membranous staining), p53 (aberrant expression indicative of mutation), HER2 score of 3 (strong complete membranous staining in >10% of cells), CD8+ and CD4+ T-cell count >0 (any detectable TILs per high-power field), EGFR (>1% membranous staining). Tumor grade G2/G3 and stage T2N0M0 were defined per standard pathological criteria. Based on the selected factors, a rating scale with different quantitative factors was constructed. The scoring scale was constructed as follows: 1 point was assigned to each patient according to the existing factors; if the data is missing, a score of 0.5 was assigned. The name of the scale cor-

Table 5. Comparison of predictive "death outcome" rating scales

Scale	Cut-off value	AuROC	Sensitivity, %	Specificity, %	Effectiveness, %	χ^2
Scale 8	3.5	0.67	76.70	49.11	62.90	46.0957
Scale 9	4.5	0.67	67.48	58.22	62.85	45.4534
Scale 10 (regression)	6.0	0.65	64.56	57.33	60.94	32.9671
Traditional scale	4.0	0.57	45.63	61.98	53.81	4.1548

Annotation: Scale 8: expression of Ck14, E-cadherin, p53 and EGFR; Expression score: FOXP3 - 0, HER2 - 3; T2N0M0; The number of CD8 cells >0 (8 factors). Scale 9: expression of Ck14, E-cadherin, p53 and EGFR; Expression score: FOXP3 - 0, HER2 - 3; T2N0M0; The number of CD8 cells >0; Degree of differentiation - G2, G3 (9 factors). Scale 10: expression of Ck14, E-cadherin, p53 and EGFR; Expression level: FOXP3 - 0, HER2 - 3; T2N0M0; The number of CD8 cells >0; Degree of differentiation - G2, G3; CD4 cell count >0 (10 factors). Traditional scale: T2N0M0; Expression score: progesterone receptor <8, estrogen receptor <8, HER2 - 3; G2.G3; Ki-67>5% (6 factors).

Table 6. Predictive scoring scale of breast cancer surrogate markers

Factor	Coefficient	Score
Expression of CK14	0.903	16
FOXP3 expression score of 0	0.833	15
T2N0M0	0.310	10
Expression of E-cadherin	0.342	6
Expression of p53	1.060	19
HER2 expression score of 3	0.709	13
CD8+ T-cell count >0	0.495	9
Expression of EGFR	0.233	4
Tumor grade G2, G3	0.296	5
CD4+ T-cell count >0	0.179	3
Total	5.607	100

responds to the number of factors covered by the scale. The selection of factors always follows the order of their influence on the expected variable "death outcome". To compare these scales as a model with the recurrence risk score scale under development, the relevant results are shown in **Table 5**.

Based on three scales (scale 8, scale 9 and scale 10), the models of "scale 8", "scale 9" and "scale 10 (regression type)" with a full score of 100 were constructed respectively. All models are compared to predict the target indicators. The results show that the proposed score regression scale has significant advantages over the traditional scale. The regression scale "regression 10" is the most effective prediction model for predicting the 10-year survival rate.

Regression prediction score scale

See **Table 6** for the regression prediction score scale. According to the scores obtained, all patients were divided into three death risk categories: low risk (score no higher than 40), medium risk (score between 40 and 60), and high risk (score higher than 60). According to the regression score scale, the risk distribution of patients in the study population is shown in Supplementary Table 1. Among the study participants, only 11.3% were identified as high-risk for malignant tumor progression.

The tumor receptor status of different predictive risk groups was analyzed based on the scores. The results showed that 85.53% of women at low risk of progression were associated with steroid status of estrogen receptor positive (ER+), progesterone receptor positive (PR+) and HER2 negative (HER2-) (see Supplementary Table 2). With the increasing risk of breast cancer progression, the proportion of this type of tumor decreased to an average of two-thirds, which were 55.89% and 55.24% in the medium risk group and high risk group, respectively.

The risk group of patients with ER-, PR-, HER2- status is shown in Supplementary Table 3. In the study group, nearly half of the patients (43.9%) were judged to have moderate risk of adverse events. Only one fifth of women (17.3%) were confirmed to have a high risk of breast cancer progression, and the proportion of low-risk patients was 38.8%.

According to the regression score scale, the 5-year and 10-year mortality assessment results of different risk groups are shown in Supplementary Table 4. Based on the statistical analysis results presented, the 5-year mortality rate in the high-risk group was only 5.8%. During the 10-year observation period, the mortality in the low-risk group increased three

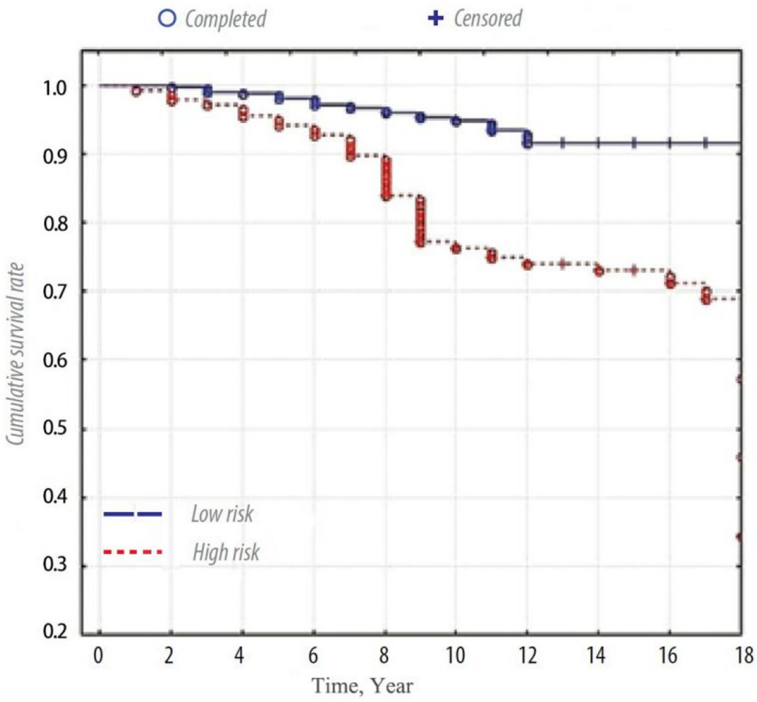

Figure 1. Kaplan Meier curve reflecting the overall survival rate of high-risk and low-risk groups in the regression score scale.

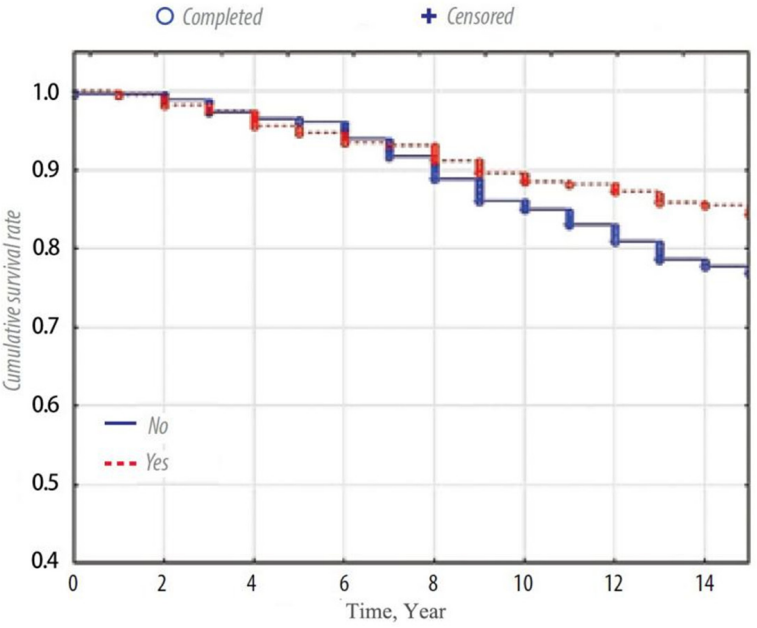

Figure 2. Kaplan Meier overall survival curve of patients in the high-risk group combined with the impact of adjuvant chemotherapy according to the regression score scale.

times compared with that in the 5-year period, and the mortality in the medium-risk group increased more than four times compared with that in the 5-year period. In the high-risk group, the 10-year mortality rate reached 30.4%, which was five times that of the 5-year mortality rate. It can be seen that there is a significant difference between the low-risk group and the medium risk group.

The survival analysis of the high-risk group and the low-risk group showed that there was a significant statistical difference between the 5-year and 10-year observation periods (P<0.00001) (see **Figure 1**). The cumulative 10-year survival rates of the high-risk group and low-risk group were 77.3% and 94.0%, respectively. In the Cox regression model, the hazard ratio of the high-risk group was 3.29 (95% CI 2.01-5.38) compared with the low-risk group. When evaluating the effect of adjuvant chemotherapy in the high-risk death risk group, the results showed that there was a statistically significant difference in the survival rate of patients receiving chemotherapy (P=0.0057) (see **Figure 2**). The 10-year cumulative survival rate was 88.6% in the high-risk group receiving chemotherapy, and 84.9% in the non-chemotherapy group. In the Cox regression model, the risk ratio of the non-chemotherapy group and chemotherapy group was 1.53 (95% CI 1.12-2.08). The 10-year evaluation results of the regression score scale for death outcomes showed sufficient sensitivity (58.05%), specificity (69.47%) and effectiveness (63.76%) (see **Table 7** and **Figure 3**). Therefore, the regression score scale is a more accurate immunohistochemical evaluation method.

Discussion

Breast cancer is the most common malignant tumor in women, with more than 2 million new cases worldwide every year [15]. In the past

Table 7. Comparison of death outcomes predicted by traditional scale and regression scale in 10-year period

Parameter	Traditional scale	Regression scale
Sensitivity, %	80.49	58.05
Specificity, %	36.92	69.47
Effectiveness, %	58.7	63.76
AuROC	0.61	0.67

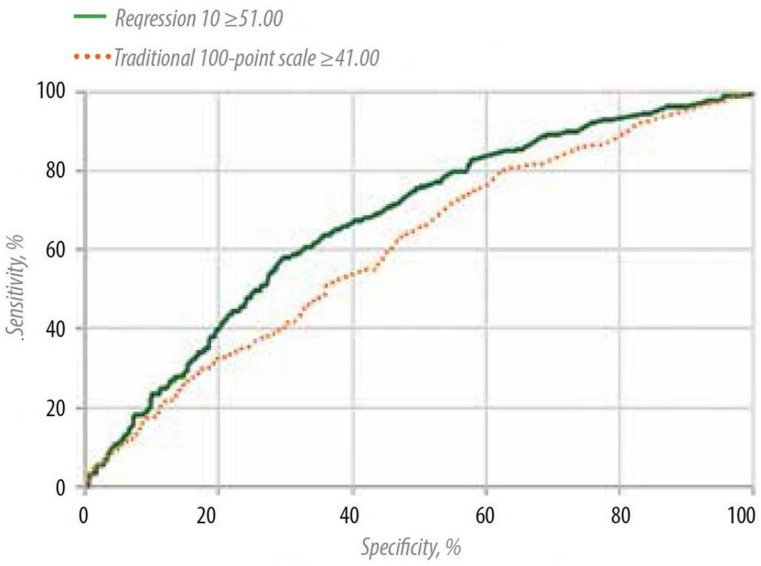

Figure 3. ROC curve of 10-year survival rate of traditional scale and regression scale.

decades, modern adjuvant therapy has undergone significant changes compared with traditional methods. Adjuvant chemotherapy, which was once regarded as a necessary means in specific clinical situations, is no longer a generally accepted treatment guideline [16].

Biomarkers play a key role in the treatment of breast cancer. Biomarkers can be used as parameters to reflect normal biological process, pathological process or treatment response [17]. Based on the latest gene research results, it has been confirmed that breast cancer shows significant genetic heterogeneity [18]. This discovery enables us to systematically classify the molecular subtypes of breast cancer according to different risk factors, morphological characteristics, treatment response characteristics and long-term prognosis, so as to formulate more accurate treatment strategies for each subtype. Although tumor molecular genetic analysis technology can deeply study tumor characteristics, however, due to the high cost of detection and the lack of prospective research evidence to support its predictive value for adjuvant therapy selection, the application of molecular genetic spectrum analysis in clinical routine is still limited. This causes some breast cancer patients to receive excessive chemotherapy.

At present, tumor molecular genetic analysis technology can carry out a detailed study of tumor characteristics, and gene expression evaluation can screen out subgroups with good prognosis and no need of adjuvant chemotherapy in patients with estrogen receptor (ER) positive/human epidermal growth factor receptor 2 (HER2) negative breast cancer. The proportion of such patients with good prognosis can exceed 40%. However, in view of the high cost of detection and the lack of prospective research evidence to support its predictive value for adjuvant therapy selection, molecular genetic spectrum analysis still has limitations in clinical routine application. This situation causes some ER positive/her2 negative breast cancer patients to suffer from excessive chemotherapy. Studies have confirmed that the results of tumor classification using alternative markers are consistent with the results of gene spectrum analysis [19].

In this study, the traditional biomarkers and new biomarkers were combined to construct the regression prediction score scale. The regression prediction score scale covers markers with little research, giving rise to high predictive value. Foxp3 is a protein involved in immune response, which plays a regulatory role as a transcription factor, and it has high prognostic value in breast cancer [20]. CK14 belongs to the cytokeratin family and is a tissue-specific protein of intermediate fibers [21]. E-cadherin is a marker of epithelial cell adhesion, and its loss of expression in cancer cells indicates partial loss of epithelial phenotype

[22]. P53 is a transcription factor that regulates the cell cycle and plays a role as an inhibitor of malignant tumor formation [23]. CD8+ T lymphocytes (cytotoxic) and CD4+ T lymphocytes (regulatory) are key cells in the immune response of breast cancer patients, and their predictive and prognostic value has also been confirmed by a number of studies [24].

In this study, survival analysis was carried out for patients in high-risk and low-risk groups, and the results showed that the difference between groups was statistically significant ($P<0.00001$). The cumulative 10-year survival rate was 77.3% in the high-risk group and 94.0% in the low-risk group. In a Cox regression model, the hazard ratio of the high-risk group and low-risk group was 3.29 (95% confidence interval 2.01-5.38). According to the analysis of the effect of adjuvant chemotherapy in patients with medium and high risk of death, the results showed that adjuvant chemotherapy had a significant advantage in the survival of patients with medium and high risk ($P=0.0057$). The cumulative 10-year survival rate was 88.6% in the chemotherapy group and 84.9% in the non-chemotherapy group. In the Cox regression model, the risk ratio of the non-chemotherapy group and chemotherapy group was 1.53 (95% confidence interval 1.12-2.08).

The multivariate prediction method in the regression score scale is helpful to improve the accuracy and reliability of prognosis assessment of breast cancer. This study focused on T1-2N0M0 breast cancer patients, constructed a prognosis scoring system based on immunohistochemical characteristics, and innovatively integrated traditional markers (estrogen receptor/progesterone receptor/human epidermal growth factor receptor 2/nuclear proliferation antigen) and new markers (cytokeratin 14/forkhead box protein P3/E-cadherin, etc.), for a total of 16 immunohistochemical indicators. Ten independent prognostic factors were screened out by logistic regression analysis, and then the scoring model was constructed. The key prognostic factors included: cytokeratin 14 expression (16 points), tumor suppressor gene p53 mutation (19 points), human epidermal growth factor receptor 23+ (13 points), CD8+ T cell infiltration (9 points). According to the scores of these 10 independent prognostic factors, the patients were stratified. Compared with the traditional scale (area under the curve is 0.61), the new model has significant advantages and can improve the accuracy of prognosis evaluation. At the same time, the cost of detection can be reduced by replacing gene detection with immunohistochemical detection. In addition, the model can effectively identify the low-risk population without adjuvant chemotherapy (accounting for 38.8%).

The prognostic landscape for early-stage breast cancer is populated by several well-established tools, each with distinct strengths and limitations. Our IHC-based regression scale positions itself within this ecosystem by offering a unique balance of comprehensiveness, accessibility, and biological insight. The classic TNM staging system provides a fundamental anatomical framework but lacks the molecular granularity needed for personalized therapy decisions in T1-2N0M0 disease, a gap our model aims to fill by integrating tumor biology and immune context.

Compared to the multi-gene assays, our model presents a complementary approach. The 21-gene Oncotype DX Recurrence Score is a validated standard for predicting chemotherapy benefit in ER+/HER2- disease, with reported sensitivity around 90% and an AuROC of approximately 0.75 in its primary validation cohorts [25]. While our scale's overall sensitivity (58.05%) is lower than that reported for Oncotype DX in its primary cohort, it offers several distinct advantages: 1) Broader applicability: It is not restricted to the ER+/HER2- subtype and showed utility in stratifying risk even within the triple-negative cohort (**Figure 3**), potentially identifying a low-risk subgroup (38.8%) that might be spared aggressive chemotherapy. 2) Biological insights: By directly quantifying tumor-infiltrating lymphocytes (CD4+/CD8+) and basal markers (CK14), our model captures elements of the tumor immune microenvironment and intrinsic biology that are not directly reported by the 21-gene assay. 3) Cost-effectiveness and accessibility: The reliance on IHC makes it a significantly more affordable and accessible option. As detailed in Supplementary Table 5, the estimated cost per test for our model is $100-$200, compared to $3,000-$4,500 for commercial genomic assays. It uses standard FFPE tissue and has a turnaround time of 2-3 days,

making it feasible for resource-constrained settings where genomic testing is not readily available or reimbursed.

Similarly, the 70-gene MammaPrint profile is a powerful tool for identifying patients with a genomically low risk of distant recurrence, with an AuROC of 0.65-0.70 [26]. Our scale's performance (AuROC 0.67) is modest compared to MammaPrint's high prognostic accuracy, but it avoids the requirement for fresh-frozen tissue and complex genomic infrastructure. The PAM50 intrinsic subtyping assay provides a robust biological classification that informs prognosis and treatment. Our model can be viewed as a proxy that approximates this biological complexity through IHC surrogates, offering a path to intrinsic subtyping in laboratories without molecular capabilities.

In conclusion, while our regression scale may not surpass the prognostic precision of advanced genomic tests in their intended populations, it serves as a pragmatic and informative alternative. Its value lies in synthesizing a wider array of biological information - including the critically relevant immune response - into a single, cost-effective score. This makes it a potentially valuable tool for risk stratification in diverse healthcare environments and for generating hypotheses about the role of the tumor microenvironment in treatment response.

This study innovatively incorporated tumor infiltrating lymphocytes (TILs, CD4+/CD8+) and other immune microenvironment indicators into the prognosis model, providing a cost-effective decision-making tool for the individualized treatment of breast cancer. A limitation of this study is the use of the entire cohort for model development without an internal validation split, which was due to the moderate sample size and the need to maximize statistical power. Therefore, future prospective studies in larger, multi-center cohorts are needed to verify the clinical applicability of the model.

Conclusion

In summary, we have developed and validated a novel IHC-based prognostic scoring scale for T1-2N0M0 breast cancer that integrates both traditional biomarkers and key elements of the tumor immune microenvironment. This scale effectively stratifies patients into distinct risk groups with significant differences in long-term survival and demonstrates potential utility in guiding adjuvant chemotherapy decisions. It offers a cost-effective and biologically comprehensive alternative to existing genomic tools, with particular relevance for resource-limited settings. However, prospective validation in larger, multi-center cohorts is imperative to confirm its clinical utility and generalizability before it can be recommended for routine clinical practice.

Acknowledgements

We express our gratitude to The First Affiliated Hospital of Gannan Medical University for support.

Written informed consent was obtained from the patient for publication of this case report and any accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Zhiyong Liu, Breast Diagnosis and Treatment Center, Huangjin Branch, The First Affiliated Hospital of Gannan Medical University, No. 128 Jinling Road, Huangjin Technology Development Zone, Ganzhou 341000, Jiangxi, China. Tel: +86-15807073854; E-mail: barton123321@163.com

References

[1] Cai Y, Dai F, Ye Y and Qian J. The global burden of breast cancer among women of reproductive age: a comprehensive analysis. Sci Rep 2025; 15: 9347.

[2] Zhai J, Wu Y, Ma F, Kaklamani V and Xu B. Advances in medical treatment of breast cancer in 2022. Cancer Innov 2023; 2: 1-17.

[3] Jose J, Hooper JD and Souza-Fonseca-Guimaraes F. Highlights of 2024. Broadening anticancer immunotherapy modalities with antibody-drug conjugates: emerging insights from clinical studies. Immunol Cell Biol 2025; 103: 530-534.

[4] Huang G, Yu Y, Su H, Gan H and Chu L. Integrating RNA-seq and scRNA-seq to explore the prognostic features and immune landscape of exosome-related genes in breast cancer metastasis. Ann Med 2025; 57: 2447917.

[5] Anand V, El-Dana F, Baran N, Borgman J, Yin Z, Zhao H, Wong ST, Andreeff M and Battula VL. GD3 synthase drives resistance to p53-induced apoptosis in breast cancer by modulating mitochondrial function. Oncogene 2025; 44: 2646-2661.
[6] Yasin R, Zafar G, Rooman Ali Syed F, Afzal S, Fatima M, Rathore Z, Chughtai A and Chughtai A. CK5/6 expression in molecular subtypes of invasive ductal carcinoma. Cureus 2024; 16: e72608.
[7] Zhang Q, Wang D, Zhuo G, Wang S, Yuan Y, Wang L, Ji L, Wan Y, Liu G and Pan Y. Intratumoral stenotrophomonas maltophilia in breast cancer: unraveling the interplay with hormone receptors and impact on tumor immunity. Int J Biol Sci 2025; 21: 974-988.
[8] Ren Z, Dharmaratne M, Liang H, Benard O, Morales-Gallego M, Suyama K, Kumar V, Fard AT, Kulkarni AS, Prystowsky M, Mar JC, Norton L and Hazan RB. Redox signalling regulates breast cancer metastasis via phenotypic and metabolic reprogramming due to p63 activation by HIF1α. Br J Cancer 2024; 130: 908-924.
[9] Jamshiya P, Ravi S, Hanuman SB, Jinkala SR, Jain A and Penumadu P. Analysis of tumor proliferation markers in early-stage luminal breast cancer: a comprehensive study using mitotic activity index, Ki-67, and phosphohistone H3 expression. Int J Surg Pathol 2025; 33: 882-890.
[10] Rätze MA, Enserink LN, Ishiyama N, van Kempen S, Veltman CH, Nijman IJ, Haakma WE, Caldas C, Bernards R, van Diest PJ, Christgen M, Koorman T and Derksen PW. Afadin loss induces breast cancer metastasis through destabilisation of E-cadherin to F-actin linkage. J Pathol 2025; 266: 26-39.
[11] Srour AM, El-Bayaa MN, Temirak A, Alanzy AL, Awad HM, Saleh A, Saleh MG and El-Sayed WA. New benzimidazole-triazole glycoconjugates as anti-cancer agents and EGFR inhibitors. Sci Rep 2025; 15: 25514.
[12] Goglia AG, Alshalalfa M, Khan A, Isakov DR, Hougen HY, Swami N, Kannikal J, Mcbride SM, Gomez DR, Punnen S, Nguyen PL, Iyengar P, Antonarakis ES, Mahal BA and Dee EC. Pan-cancer genomic analysis reveals FOXA1 amplification is associated with adverse outcomes in non-small cell lung, prostate, and breast cancers. J Natl Cancer Inst 2025; 117: 188-197.
[13] Omar M, Harrell JC, Tamimi R, Marchionni L, Erdogan C, Nakshatri H and Ince TA. A triple hormone receptor ER, AR, and VDR signature is a robust prognosis predictor in breast cancer. Breast Cancer Res 2024; 26: 132.
[14] Du T, Yuan Y, Sun S, Gao Z and Li X. Integrating traditional biomarkers and emerging predictors to assess neoadjuvant chemotherapy efficacy in breast cancer: a multifactorial analysis of Ki-67, CDK4, EGFR, TILs and ctDNA. BMC Womens Health 2024; 24: 674.
[15] Liu L, Zhou P, Hou L, Kao C, Zhang Z, Wang D, Yu L, Wang F, Wang Y and Yu Z. Development and performance of female breast cancer incidence risk prediction models: a systematic review and meta-analysis. Ann Med 2025; 57: 2534522.
[16] Bhimani J, Wang P, Gallagher GB, O'Connell K, Blinder V, Burganowski R, Ergas IJ, Griggs JJ, Heon N, Kolevska T, Kotsurovskyy Y, Kroenke CH, Laurent CA, Liu R, Nakata KG, Persaud S, Roh JM, Tabatabai S, Valice E, Bandera EV, Bowles EJA, Kushi LH and Kantor ED. Patient factors and modifications to intended chemotherapy for women with Stages I-IIIA breast cancer. Int J Cancer 2025; 157: 1342-1353.
[17] McClurg DP, Wong M, Luo-Yng Tay K, Urquhart G, Soh FY, Saad Abdalla Al-Zawi A, Masannat Y, Elsberger B and Speirs V. The prognostic significance of the serum inflammatory marker ratios, neutrophil-lymphocyte, platelet-lymphocyte and monocyte-lymphocyte in male breast cancer. Cancer Invest 2025; 43: 391-398.
[18] Enoma D. Genomics in clinical trials for breast cancer. Brief Funct Genomics 2024; 23: 325-334.
[19] Jiang Z, Xu B, Sun B, Yang B, Lu S, Li M, Zhang J, Qi L and Wu Q. Germline variants analysis of Chinese breast cancer patients reveals numerous alterations in homologous recombination genes. Future Sci OA 2025; 11: 2458432.
[20] Li L, Zhang Z, Huang N, Ren J, Qin Y and Luo Y. IGF1R activates FOXP3-β-catenin signaling to promote breast cancer development. Breast Cancer Res Treat 2025; 211: 467-478.
[21] Sousa B, Paredes J, Milanezi F, Lopes N, Martins D, Dufloth R, Vieira D, Albergaria A, Veronese L, Carneiro V, Carvalho S, Costa JL, Zeferino L and Schmitt F. P-cadherin, vimentin and CK14 for identification of basal-like phenotype in breast carcinomas: an immunohistochemical study. Histol Histopathol 2010; 25: 963-974.
[22] Ashrafi P, Sari S, Javani Jouni F, Zafari J and Asgari F. Potentiated effects of photobiomodulation and celecoxib on the epithelial-mesenchymal transition signaling of E-Cadherin, N-Cadherin, α-SMA in breast cancer cells, MCF7, and MDA-MB-231. Photobiomodul Photomed Laser Surg 2025; 43: 115-123.
[23] Kim M, Lee M, Lee A, Choi BO, Park WC, Kim SH, Lee J and Kang J. Correlating p53 immunostaining patterns with somatic TP53 muta-

tion and functional properties of mutant p53 in triple-negative breast cancer. Histopathology 2025; 87: 299-309.
[24] Jagtap SV. Evaluation of CD4+ T-cells and CD8+ T-cells in triple-negative invasive breast cancer. Indian J Pathol Microbiol 2018; 61: 477-478.
[25] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J and Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817-2826.
[26] Cardoso F, van't Veer LJ, Bogaerts J, Slaets L, Viale G, Delaloge S, Pierga JY, Brain E, Causeret S, DeLorenzi M, Glas AM, Golfinopoulos V, Goulioti T, Knox S, Matos E, Meulemans B, Neijenhuis PA, Nitz U, Passalacqua R, Ravdin P, Rubio IT, Saghatchian M, Smilde TJ, Sotiriou C, Stork L, Straehle C, Thomas G, Thompson AM, van der Hoeven JM, Vuylsteke P, Bernards R, Tryfonidis K, Rutgers E and Piccart M; MINDACT Investigators. 70-gene signature as an aid to treatment decisions in early-stage breast cancer. N Engl J Med 2016; 375: 717-729.

Supplementary Table 1. Patient risk grouping based on regression score scale

Low risk	27.0%
Intermediate risk	61.7%
High risk	11.3%

Supplementary Table 2. Assignment of patients with different receptor status to death risk subgroups according to the regression score scale

	ER+, PR+, HER2-	HER2+	ER-, PR-, HER2-
Low risk	85.83%	0.57%	13.60%
Intermediate risk	55.89%	17.11%	27.00%
High risk	55.24%	19.05%	25.71%

Supplementary Table 3. Risk stratification of patients with ER-, PR-, HER2- status according to the regression score scale (formerly **Figure 3**)

Low risk	38.8%
Intermediate risk	43.9%
High risk	17.3%

Supplementary Table 4. Mortality in the low-, medium-, and high-risk groups as assessed by the regression score scale

	5 years	10 years
Low risk	2.5%	7.7%
Intermediate risk	4.2%	19.5%
High risk	5.8%	30.4%

Supplementary Table 5. Comparison of the IHC-based prognostic scale with commercial genomic assays (Cost, Logistics, and Features)

Feature	Our IHC-Based Model	Oncotype DX® (21-gene)	MammaPrint® (70-gene)	Prosigna® (PAM50)
Technology Platform	Immunohistochemistry (IHC)	RT-PCR from FFPE	DNA Microarray from Fresh-Frozen or FFPE	RT-PCR from FFPE
Key Biological Insights	Hormone receptors, HER2, proliferation (Ki-67), Tumor Immune Microenvironment (CD4/CD8), basal markers (CK14), etc.	Proliferation, ER, HER2, invasion genes	proliferation, invasion, metastasis, stromal integrity genes	Intrinsic molecular subtyping (Luminal A/B, HER2-enriched, Basal-like)
Primary Clinical Indication	Prognosis & risk stratification for all T1-2N0M0 subtypes	Predicting chemotherapy benefit in ER+/HER2-, LN-	Predicting distant recurrence risk in early-stage cancer	Prognostic risk category & intrinsic subtyping
Tissue Requirements	Standard FFPE tissue block	FFPE tissue block	Fresh-frozen tissue *or* specially processed FFPE	FFPE tissue block
Estimated Cost per Test	$100-$200 (Estimated reagent and labor cost)	$3,000-$4,000	$3,000-$4,200	$3,500-$4,500
Typical Turnaround Time	2-3 days	10-14 days	10-14 days (longer for international sites)	10-14 days
Accessibility in Resource-Limited Settings	High (Uses standard pathology lab equipment and expertise)	Low (Centralized labs, complex logistics, high cost)	Very Low (Fresh-frozen requirement is a major barrier)	Low (Centralized testing model)
Applicability to TNBC Subtype	Yes (Shown to stratify risk in TNBC cohort, **Figure 3**)	No (Not validated for TNBC)	Yes	Yes