Breast Cancer: Targets and Therapy

Dovepress

open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Elevated Risk of Adverse Prognosis in Patients with T2-3 Stage Breast Cancer Exhibiting Non-Pathological Complete Response Following Neoadjuvant Chemotherapy: Significance of Regenerating Islet-Derived Family Member 4

Fan Li[1,*], Chuan-Guo Chen[2,*], Jiao-Fei Wei[1,3], Jia-Wen Lin[4], Zi-Ang Dou[3,4], Jun Shen[1], Shu-Qin Li[1,3,4]

[1]Department of Breast Surgery, the First People's Hospital of Lianyungang, The Affiliated Hospital of XuZhou Medical University, Lianyungang, Jiangsu, People's Republic of China; [2]Department of General Surgery, Nanjing Meishan Hospital, Nanjing, Jiangsu, People's Republic of China; [3]Department of Clinical Medicine, Jinzhou Medical University, Jinzhou, Liaoning, People's Republic of China; [4]Department of clinical medicine, Lianyungang Clinical College of Nanjing Medical University, Lianyungang, Jiangsu, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jun Shen; Shu-Qin Li, Department of Breast Surgery, the First People's Hospital of Lianyungang, The Affiliated Hospital of XuZhou Medical University, No. 6 Zhenhua East Road, High-tech Square, Lianyungang, Jiangsu Province, 222002, People's Republic of China, Tel +8618961325320; +8618961325781, Email shenjun2257@126.com; shuqinlilsh@126.com

Objective: In this study, we aimed to establish the role of regenerating islet-derived family member 4 (Reg IV) as an independent risk factor and prognostic predictor in patients with T2-3 stage breast cancer who exhibit a non-pathological complete response (non-pCR) following neoadjuvant chemotherapy (NACT). Additionally, we examined the potential correlation and interaction between Reg IV and epidermal growth factor receptor (EGFR).

Methods: A total of 67 patients with T2-3 stage breast cancer exhibiting non-pCR after NACT between September 2019 and December 2021 were included in this study. The analysis involved Kaplan–Meier survival comparisons, pooled hazard ratios for risk quantification, Cox regression analysis to isolate the impact of Reg IV on prognosis, Riskplots for visualizing risk profiles, and SHAP analysis to assess the importance of variables in predicting outcomes.

Results: The findings indicate that patients positive for Reg IV had a significantly poorer prognosis (HR: 2.62, 95% CI: 1.06–6.47). Co-expression of Reg IV and EGFR was associated with the worst outcomes compared to patients negative for both markers. Cox regression analysis confirmed the independent prognostic impact of Reg IV (HR: 2.63, 95% CI: 1.66–3.59). Riskplot analysis showed that patients positive for both Reg IV and EGFR predominantly experienced disease progression. SHAP analysis further reinforced the significant effect of Reg IV on the disease course, without substantial interaction with EGFR.

Conclusion: Reg IV may serve as an independent risk factor and predictive marker for adverse outcomes in patients with T2-3 stage breast cancer who do not achieve non-pCR following NACT.

Keywords: breast cancer, NACT, neoadjuvant chemotherapy, non-pCR, non-pathologic complete response, reg IV, risk factor

Introduction

Regenerating islet-derived family member 4 (Reg IV), a member of the calcium-dependent lectin superfamily, is implicated in enhancing the invasive and migratory capabilities of tumor cells.[1] Previous research has identified Reg IV as a potential factor closely associated with adverse prognostic indicators and clinicopathological characteristics in various cancer types, including colon, rectal, gastric, gallbladder, pancreatic, ovarian, prostate, and lung cancers.[2] Neoadjuvant chemotherapy (NACT), which reduces the risk of recurrence, is now the standard treatment for breast

Received: 15 April 2024
Accepted: 31 August 2024
Published: 12 September 2024

cancer and is widely utilized in clinical practice.[3] Patients who achieve a pathological complete response (pCR) tend to have a better prognosis than those who do not (non-pCR), highlighting the need for targeted treatment strategies for non-pCR patients.[4] Optimizing treatment for non-pCR patients is crucial for further advancements in breast cancer outcomes.

Although the biological function of Reg IV remains inadequately understood, studies have demonstrated its co-expression with epidermal growth factor receptor (EGFR) in gastric and colon cancers.[5,6] Reg IV is known to activate EGFR, thereby promoting cancer cell growth.[7] However, there is limited research on the role of Reg IV in breast tumor progression, particularly regarding its prognostic significance in patients undergoing NACT who do not achieve non-pCR status.[3,4,8] The objective of this study is to elucidate the involvement of Reg IV in the progression of T2-3 stage breast cancer in patients treated with NACT who fail to attain non-pCR status.

Materials and Methods

Study Participants

The study included 67 patients with breast cancer who underwent NACT at the First People's Hospital of Lianyungang between September 2019 and December 2021. None of the participants had received radiotherapy or tumor endocrine therapy prior to NACT. Imaging examinations at the initial visit confirmed the absence of distant metastases or other malignant tumors. All participants completed the planned NACT, followed by surgical intervention and pathological examinations. Medical data was collected, and all participants voluntarily participated in the study. This study was approved by the Ethics Committee of the First People's Hospital of Lianyungang (Approval No.: KY-20230410002-01). Informed consent was obtained from all study participants.

Observation Indicators

Patient characteristics and pathological data were extracted from the medical record systems and pathological report systems of the hospital, respectively. Specimens were stored in the specimen repository of the Pathology Department. Specimens were fixed in a 3.7% neutral formaldehyde solution, followed by routine dehydration, clearing, paraffin infiltration, and sectioning at a thickness of 4 μm. The sections underwent EnVision Two-Step immunohistochemical staining and were examined under a light microscope. Paraffin sections were baked at 60 °C for 90 minutes, dewaxed using xylene, hydrated using gradient ethanol, and blocked for endogenous peroxidase activity with 3% hydrogen peroxide. Antigen retrieval for Reg IV and EGFR was performed using high-pressure citrate buffer and ethylenediaminetetraacetic acid (EDTA), respectively. Sections were then blocked with goat serum for 20 minutes and incubated with primary antibodies overnight at 4 °C (Reg IV at a 1:100 dilution; EGFR in working solution). Secondary antibodies were incubated in a constant temperature box at 37 °C for 40 minutes. Staining was conducted using 3,3′-diaminobenzidine (DAB) and counterstained with hematoxylin. Phosphate-buffered saline (PBS) served as the control for the primary antibody. Specimens were scored based on staining intensity and the percentage of positive cells. Five to ten high-power fields were randomly selected to count positive cells, with the average percentage calculated. Scoring gradients for the percentage of positive cells were: 0 points (no positive cells), 1 point (1% to 25%), 2 points (26% to 50%), 3 points (51% to 75%), and 4 points (76% to 100%). Staining intensity was scored as: 0 points (no staining), 1 point (pale yellow staining), 2 points (brownish yellow staining), and 3 points (brown staining). The combined score of staining intensity and percentage of positive cells was classified as negative (≤ 3 points) or positive (> 3 points). Results were assessed in a double-blind manner by two chief pathologists.

Breast cancer staging followed the tumor-node-metastasis (TNM) system of the Union for International Cancer Control (UICC): T1 (< 20 mm), T2 (20–50 mm), and T3 (> 50 mm).[9] Lymph node metastasis was classified as N0 (none), N1 (1–3), N2 (4–8), and N3 (≥ 9). Estrogen receptor (ER) and progesterone receptor (PR) expression in tumor cells was classified as positive at less than or equal to 1%, with low expression between 1% and 30%, and high expression above 30%. Negative expression was below 1%. Human epidermal growth factor receptor 2 (HER-2) positivity was determined in accordance with the guidelines for HER2 testing in patients with breast cancer established by the American Society of Clinical Oncology (ASCO) and College of American Pathologists (CAP).[10] A positive HER2 result was characterized by an immunohistochemistry (IHC) staining of 3+ (A high expression—the presence of more

than 30% of invasive tumor cells), and fluorescence in situ hybridization (FISH) demonstrating more than 6 copies of HER2 genes per nucleus or a FISH ratio (chromosome 17/HER2 gene signal) exceeding 2.2. Negative results were indicated by an IHC staining of 0 or 1+, a FISH result revealing fewer than 4.0 HER2 gene copies per nucleus, or a FISH ratio below 1.8. HER-2 (3+), as well as HER-2 (2+) with FISH (+), were defined as HER-2 positive; while (0) (1+) (2+) with FISH (-) were defined as HER-2 negative. Ki-67 expression was classified as low (< 20%) or high (≥ 20%).

Classifications of Pathologic Complete Response

In accordance with the 8th edition guidelines of the American Joint Committee on Cancer (AJCC), pathologic complete response (pCR) is defined as the absence of invasive breast cancer in both the primary tumor site and axillary lymph nodes upon examination of the surgical specimen following NACT, denoted as ypT0/ypTis, ypN0. This definition accommodates the presence of residual non-invasive breast cancer, like ductal carcinoma in situ (DCIS).

Definitions of Disease Progression

The progression events of the disease (endpoint events) are delineated as follows:

1) Soft tissue metastasis and recurrence: Confirmed via biopsy, involving sites such as ipsilateral supraclavicular lymph nodes, cervical lymph nodes, contralateral axillary lymph nodes, supraclavicular lymph nodes, cervical lymph nodes, chest wall soft tissues, and ipsilateral axillary lymph nodes (excluding bilateral primary disease onset).

2) Liver and lung metastasis: Confirmed through biopsy under computed tomography (CT) or color Doppler ultrasound guidance.

3) Bone metastasis: Confirmed either based on pathology or determined by expert assessment using ECT in combination with X-ray, CT, magnetic resonance imaging (MRI).

4) Brain metastasis: Indicated by imaging and symptoms such as headache.

5) Bone marrow metastasis: Confirmed through bone marrow biopsy and pathology.

6) Ovarian metastasis: Diagnosed via pathological biopsy.

7) Other sites: Mediastinal lymph node, internal mammary lymph node, or retroperitoneal lymph node metastasis confirmed via biopsy or diagnosed as metastasis from breast cancer by positron emission tomography and computed tomography (PET-CT).

Statistical Analysis

Statistical analysis and visualization were conducted using R 4.0.2 software. Comparisons of clinical characteristics between groups were performed using the chi-squared test or Fisher's exact test for categorical variables. Correlation between variables was examined through cluster analysis, while the relationship between Reg IV and EGFR was analyzed using Spearman correlation analysis. Kaplan–Meier curves were generated to illustrate the expression of Reg IV and EGFR. Spearman correlation analysis was used to calculate correlations among all variables, with the correlation matrix visualized through original vector distance classification in cluster analysis.

Variations in hazard ratios (HR) across different data groups were computed using 10-fold cross-validation and aggregated via meta-analysis. Riskplot analysis was conducted to assess risk variables and compare them across various models using the area under the curve (AUC). Model stability was assessed by plotting time-dependent receiver operating characteristic (ROC) curves. The significance of each variable in the model was reflected in its contribution to the predictive outcome, quantitatively assessed and visually presented using the SHapley Additive exPlanations (SHAP) method. SHAP values provide a comprehensive analysis of the marginal contributions of each feature to the output of the model, indicating the variables exerting the greatest influence on the overall prediction results. This approach facilitates a deeper understanding of the internal mechanisms of the model.

Differences between models were assessed using the Net Reclassification Improvement (NRI) and Integrated Discrimination Improvement (IDI).[11,12] Models were depicted using a nomogram, and their stability at different time points was assessed through calibration curves. Statistical significance was defined as a P-value less than 0.05 ($P < 0.05$).

Results

Study Participants

This study included 67 patients with breast cancer who did not achieve non-pCR following NACT. Among these patients, 36 (54%) were negative for Reg IV, while 31 (46%) were positive. Additionally, 32 (48%) were negative for EGFR, while 35 (52%) were positive. According to clinical staging, 27 patients (40%) were classified as stage T2, and 40 patients (60%) as stage T3. Lymph node involvement was observed in 43 patients (64%), while 24 patients (36%) were negative for lymph nodes. The study cohort was divided into training and validation datasets in a 7:3 ratio for further analysis (Table 1). Over a median follow-up period of 21 months (range: 18–30 months), 19 patients experienced disease progression events. Among these patients, 6 (32%) were negative for Reg IV, while 13 (68%) were positive, and the difference was statistically significant. No statistically significant differences were observed in EGFR, Ki-67, ER, PR, HER-2, and other assessed indicators (Table S1). All patients followed the chemotherapy regimen and dosage outlined in the CSCO guidelines, with 76% (56/67) receiving postoperative adjuvant radiotherapy. Details of the specific chemotherapy and radiotherapy regimens are provided in Table S2.

Table 1 Baseline Indexes

Characteristic	All Data Set	Train Data Set	Test Data Set
	N = 67[a]	N = 50[a]	N = 17[a]
Treatment			
AC	3(4.5%)	3(6.0%)	0(0%)
AC-T	15(22%)	11(22%)	4(24%)
AC-THP	1(1.5%)	1(2.0%)	0(0%)
TAC	20(30%)	16(32%)	4(24%)
TCbH	2(3.0%)	2(4.0%)	0(0%)
TCbHP	17(25%)	13(26%)	4(24%)
TCHP	1(1.5%)	1(2.0%)	0(0%)
THP	4(6.0%)	2(4.0%)	2(12%)
TP	4(6.0%)	1(2.0%)	3(18%)
Age			
<40	8(12%)	7(14%)	1(5.9%)
≥40	59(88%)	43(86%)	16(94%)
Type			
TNBC	17(25%)	10(20%)	7(41%)
LUMI	12(18%)	10(20%)	2(12%)
HER2	38(57%)	30(60%)	8(47%)
Clinic stage			
T2	27(40%)	18(36%)	9(53%)
T3	40(60%)	32(64%)	8(47%)

(*Continued*)

Table 1 (Continued).

Characteristic	All Data Set	Train Data Set	Test Data Set
	N = 67[a]	N = 50[a]	N = 17[a]
Menstrual status			
Premenopause	25(37%)	21(42%)	4(24%)
Postmenopause	42(63%)	29(58%)	13(76%)
T stage			
T2	50(75%)	38(76%)	12(71%)
T3	17(25%)	12(24%)	5(29%)
N stage			
Neg	24(36%)	18(36%)	6(35%)
Pos	43(64%)	32(64%)	11(65%)
ER			
Neg	24(36%)	17(34%)	7(41%)
Pos	43(64%)	33(66%)	10(59%)
PR			
Neg	39(58%)	28(56%)	11(65%)
Pos	28(42%)	22(44%)	6(35%)
HER-2			
Neg	41(61%)	29(58%)	12(71%)
Pos	26(39%)	21(42%)	5(29%)
Ki-67			
<20	16(24%)	14(28%)	2(12%)
≥20	51(76%)	36(72%)	15(88%)
EGFR			
Neg	32(48%)	26(52%)	6(35%)
Pos	35(52%)	24(48%)	11(65%)
Reg IV			
Neg	36(54%)	26(52%)	10(59%)
Pos	31(46%)	24(48%)	7(41%)

Notes: [a]n(%); Mean(SD).

Cluster Analysis and Kaplan–Meier Survival Analysis

The cluster analysis revealed distinct clustering patterns concerning ER and PR status, treatment type, HER-2 expression, patient age, menstrual status, as well as T stage and clinical stage. Spearman's rank correlation tests, in conjunction with the cluster analysis, indicated a correlation coefficient of 0.4 between Reg IV and EGFR ($P > 0.05$), revealing no statistically significant correlation between these two variables. (Figure 1A and B). Kaplan–Meier survival analysis

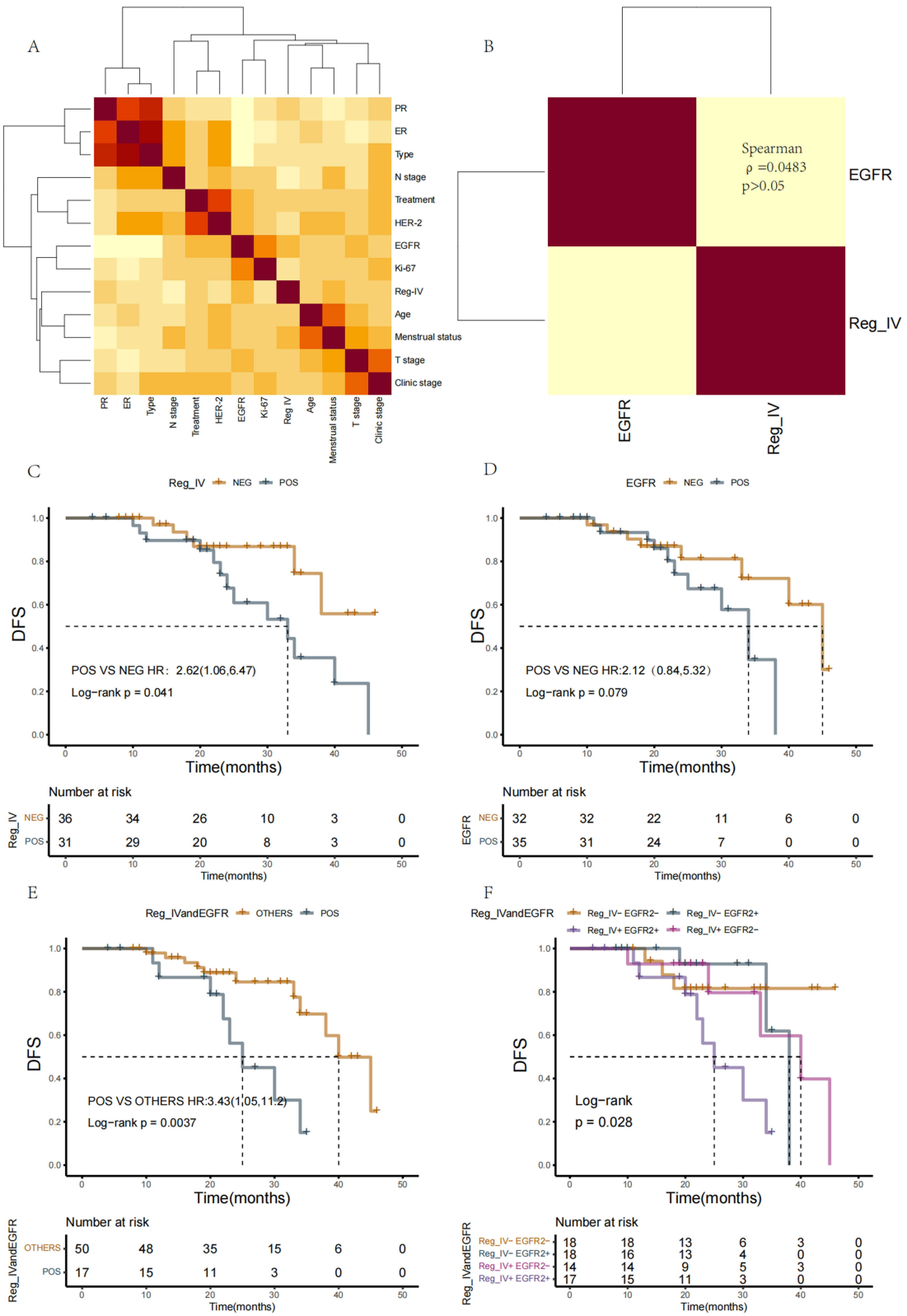

Figure 1 Cluster analysis and Kaplan–Meier survival analysis. (**A**). The cluster analysis graph illustrates notable clustering patterns among ER, PR, and type, treatment and HER-2, age and menstrual status, and T stage and clinic stage. (**B**). The analysis graph depicts the correlation between Reg IV and EGFR. (**C**). Kaplan–Meier survival analysis demonstrates the survival outcomes of patients with positive Reg IV compared to those with negative Reg IV. (**D**). Kaplan–Meier survival analysis reveals the survival outcomes of patients with positive EGFR compared to those with negative EGFR. (**E**). Kaplan–Meier survival analysis presents the survival outcomes of patients with both positive Reg IV and EGFR compared to other patients. (**F**). Kaplan–Meier survival analysis compares the survival outcomes among different groups, including Reg IV positive vs EGFR positive, Reg IV negative vs EGFR negative, Reg IV positive vs EGFR negative, and Reg IV negative vs EGFR positive.

demonstrated that patients with positive Reg IV expression had a poorer prognosis compared to those with negative Reg IV expression, with an HR of 2.62 (95% CI: 1.06, 6.47) and a log-rank *P*-value of < 0.05. Similarly, patients with positive EGFR expression had a worse prognosis compared to those with negative EGFR expression, with an HR of 2.12 (95% CI: 0.84, 5.32) and a log-rank *P*-value of < 0.05. Patients positive for both Reg IV and EGFR had the poorest prognosis among all groups, with an HR of 3.43 (95% CI: 1.05, 11.2) and a log-rank *P*-value of < 0.05. Stratified analysis indicated that patients positive for both Reg IV and EGFR had the poorest prognosis, while those negative for both exhibited the most favorable outcomes. Patients positive for Reg IV but negative for EGFR, as well as vice versa, had similar prognoses (Figure 1C–F).

Cox Regression Analysis

The correlation of each indicator with disease-free survival (DFS) was examined through single-factor Cox regression analysis and AUC calculations (Table 2 and Figure S1). Single-factor Cox regression analysis identified Reg IV as an independent prognostic factor, exhibiting an HR of 2.63 (95% CI: 1.66, 3.59) ($P < 0.05$), along with an AUC of 0.635 (95% CI: 0.527, 0.782). The EGFR value served as the threshold for distinguishing between positive and negative results, with an HR of 2.45 (95% CI: 1.43, 3.466), $P = 0.084$, and an AUC of 0.539 (95% CI: 0.405, 0.674). Other indicators, such as T stage, N stage, ER, and PR, showed no statistically significant impact on DFS. A 10-fold cross-validation was conducted to validate the influence of Reg IV on prognosis, revealing HR fluctuations across different datasets, with minimum and maximum HR values of 1.94 (95% CI: 1.08, 3.47) and 3.82 (95% CI: 2.82, 5.19), respectively. The HRs were aggregated through meta-analysis, with pooled HRs estimated at 2.92 (95% CI: 2.57, 3.31) and 2.78 (95% CI: 2.35, 3.30) for the random-effects and fixed-effects models, respectively, and an I2 value of 39%. These findings indicate that elevated expression of Reg IV is associated with a poor prognosis (Figure 2A).

Riskplot ROC and Time-Dependent ROC

Riskplot analysis revealed a gradual increase in estimated risk among different disease progression events, with a cutoff value of 0.08 determined through AUC calculations. Among patients experiencing disease progression events, those positive for both Reg IV and EGFR constituted the largest proportion, while patients positive for ER and PR had a comparatively better prognosis. No discernible discriminatory patterns were observed among other indicators like T stage and N stage (Figure 2B). The entire dataset was used to assess the influences on DFS from Reg IV alone, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging. The AUC values for these models were 0.655 (95% CI: 0.527−0.782), 0.669 (95% CI: 0.526−0.813), and 0.678 (95% CI: 0.532−0.824), respectively (Figure 2C–E).

Table 2 Cox Regression Analysis in Data.nonPCR

	HR	LCI	UCI	P	ROC	LAUC	UAUC
Treatment	0.9928	0.8039	1.1816	0.940153	0.559	0.395	0.723
Age	0.9651	−0.5239	2.454	0.962659	0.49	0.405	0.575
Type	0.8387	0.3481	1.3292	0.482034	0.556	0.409	0.703
Clinic stage	1.8727	0.8345	2.9109	0.236236	0.598	0.473	0.722
Menstrual status	0.8733	−0.0612	1.8079	0.77638	0.533	0.4	0.666
T stage	1.4196	0.4333	2.4058	0.486264	0.58	0.454	0.706
N stage	0.7612	−0.192	1.7143	0.574686	0.507	0.376	0.638
ER	0.5262	−0.3828	1.4352	0.166264	0.617	0.485	0.75
PR	0.6965	−0.2217	1.6147	0.440138	0.498	0.364	0.632
HER-2	0.9286	−0.1141	1.9714	0.889321	0.587	0.463	0.711
Ki-67	1.0286	0.0517	2.0054	0.954928	0.446	0.324	0.568
EGFR	2.4487	1.4314	3.466	0.084438	0.539	0.405	0.674
Reg-IV	2.6314	1.6631	3.5997	0.050186	0.655	0.527	0.782

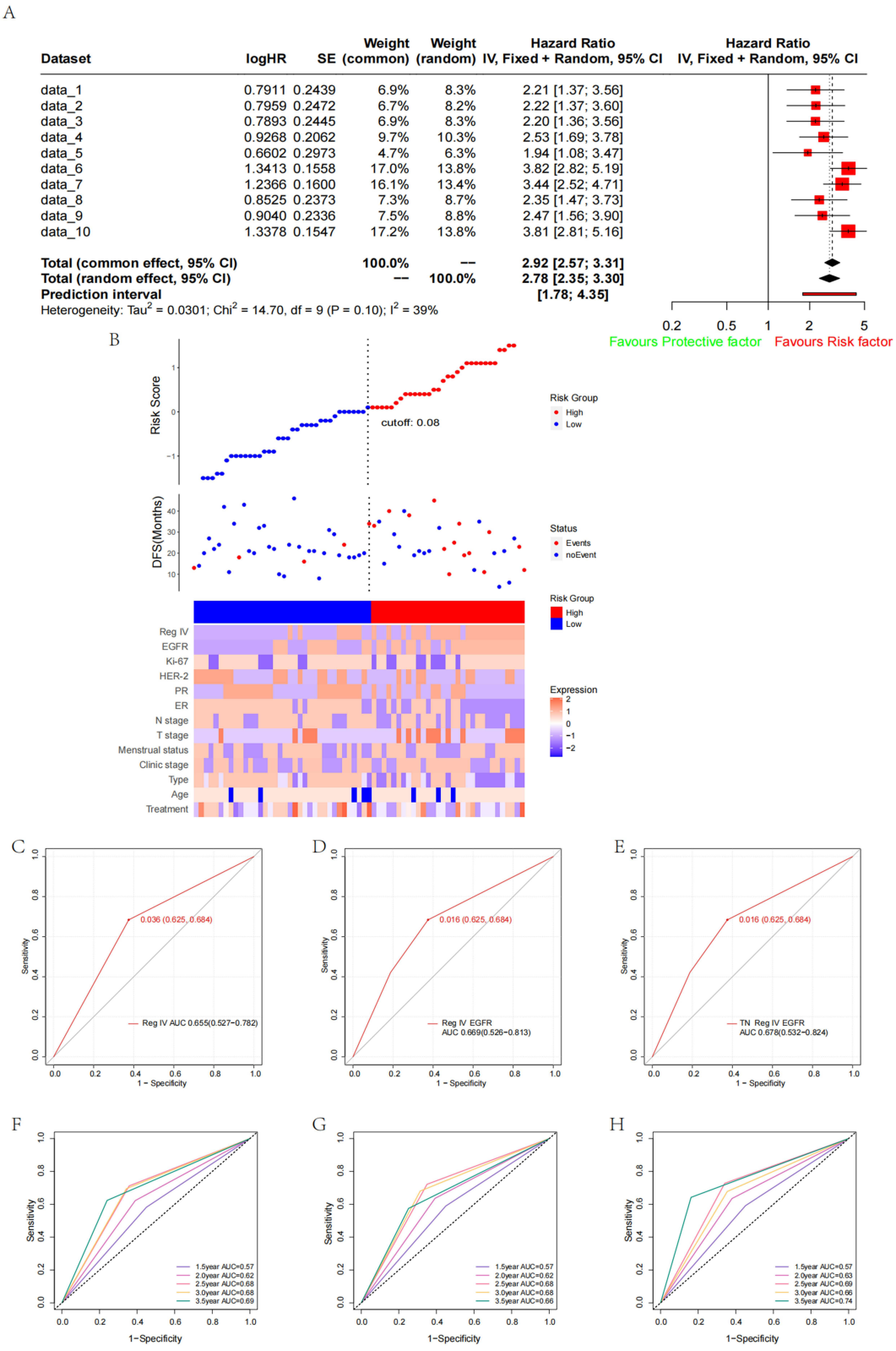

Figure 2 Cox regression analysis and the ROC curve. (**A**). The methodology used in this study involved a 10-fold cross-validation, META analysis, and the pooling of HRs. (**B**). Riskplot analysis revealed that as the assessed risks increased, there was a corresponding increase in discrimination of disease progression events. The calculation, based on the AUC curve, determined a cutoff value of 0.08. Among patients experiencing disease progression events, those with both positive Reg IV and EGFR accounted for the largest proportion. (**C**–**E**). These figures represent the ROC curve for predicting DFS using Reg IV alone, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging, respectively. (**F**–**H**). These figures represent the time-dependent ROC curve for predicting DFS using Reg IV alone, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging, respectively.

Time-dependent ROC curves showed that the Reg IV model and the combined models involving EGFR and TN staging exhibited stable efficacy across various time points. However, the addition of EGFR and TN staging did not significantly enhance performance compared to the Reg IV model alone (Figure 2F–H).

Interaction of EGFR with SHAP and Reg IV

Interpretability analysis of model variables identified Reg IV as the most explanatory variable, with statistical significance in variable analysis (Figure 3A). In the time-dependent ROC curve, Reg IV was the most crucial variable (Figure 3B). Models incorporating Reg IV alone, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging demonstrated effective discriminatory model performance. The Reg IV model displayed a favorable long-term prognosis extending over 30 months. The addition of EGFR enhanced predictive stability up to 15 months, although incorporating TN staging did not yield significant enhancement (Figure 3C–E).

Among patients with high Reg IV expression, the impact of EGFR on prognosis approached statistical significance (β = 1.169, SE = 0.682, Wald = 1.71, P = 0.086). Among patients with low Reg IV expression, EGFR exhibited no significant association with prognosis (β = 0.243, SE = 0.831, Wald = 0.29, P = 0.769). A difference was observed in the correlation between Reg IV expression and prognosis in patients with high and low EGFR expressions. Among patients with high EGFR expression, Reg IV significantly influenced prognosis (β = 1.599, SE = 0.792, Wald = 2.02, P = 0.043). However, among patients with low EGFR expression, the association between Reg IV and prognosis was not statistically significant (β = 0.648, SE = 0.731, Wald = 0.89, P = 0.375). Interaction assessment indicated no significant interaction between EGFR and Reg IV (β = 0.575, SE = 1.0, Wald = 0.58, P = 0.565).

NRI and IDI

NRI and IDI curves were computed for 24 months and 36 months to ascertain differences between the Reg IV model, the model combining Reg IV with EGFR, and the model combining Reg IV with EGFR and TN staging. For the Reg IV model versus the model combining Reg IV with EGFR, the 24-month NRI values were: NRI: 0.0217762 (−0.2777795, 0.5743983), NRI+: −0.1920152 (−0.3512066, 0.7730022), and NRI-: 0.2137914 (−0.4883770, 0.5242998). The 36-month NRI values were: NRI: 0.69712974 (−0.4133891, 0.9723828), NRI+: 0.36417170 (−0.3027230, 0.9623543), and NRI-: 0.33295804 (−0.4165496, 0.8235251). In both cases, there was no statistically significant difference (Figure 4A and B). For the Reg IV model versus the model combining Reg IV with EGFR and TN staging, the 24-month NRI values were: NRI: −0.009646097 (−0.2592458, 0.7961692), NRI+: −0.067205325 (−0.4099516, 0.6418267), and NRI-: 0.057559228 (−0.2403363, 0.5313693). The 36-month NRI values were: NRI: 0.68009728 (−0.3919524, 1.1538152), NRI+: 0.39968562 (−0.3919524, 1.0283311), and NRI-: 0.28041166 (−0.3730602, 0.7421165), again with no statistically significant difference (Figure 4C and D). IDI comparisons for 24 months between the Reg IV model and the model combining Reg IV with EGFR showed an IDI of 0.058 (−0.005, 0.198) (P = 0.109), and for 36 months, it was 0.155 (−0.040, 0.426) (P = 0.149) (Figure 4E and F). For the Reg IV model versus the model combining Reg IV with EGFR and TN staging, the 24-month IDI was 0.087 (0.005,0.324) (P = 0.03), and the 36-month IDI was 0.131 (−0.091,0.444) (P = 0.209) (Figure 4G and H).

ROC Curves and Calibration Curves for the Training and Validation Sets

To validate the stability of the Reg IV model in prognostic prediction, the dataset was partitioned into the training and validation sets. The ROC curves for the training and validation sets were calculated as 0.671 (95% CI: 0.512–0.830) and 0.584 (95% CI: 0.347–0.822), respectively (Figure 5A and B). Calibration curves for 24 and 36 months demonstrated satisfactory performance in the training sets. However, the calibration curve for 24 months indicated poor performance in the validation set, while the calibration curve for 36 months revealed good performance in the validation set. These results suggest that incorporating TN staging and EGFR may enhance efficiency due to easier access to clinical data, however Reg IV remains a reliable predictive factor (Figures 5C–F). The prediction efficiency for the 24-month outcome in the verification set model was unstable, primarily due to the limited number of patients with outcomes at 24 months. Despite a similar proportion of outcome events during randomization, the verification set had too few outcome events at

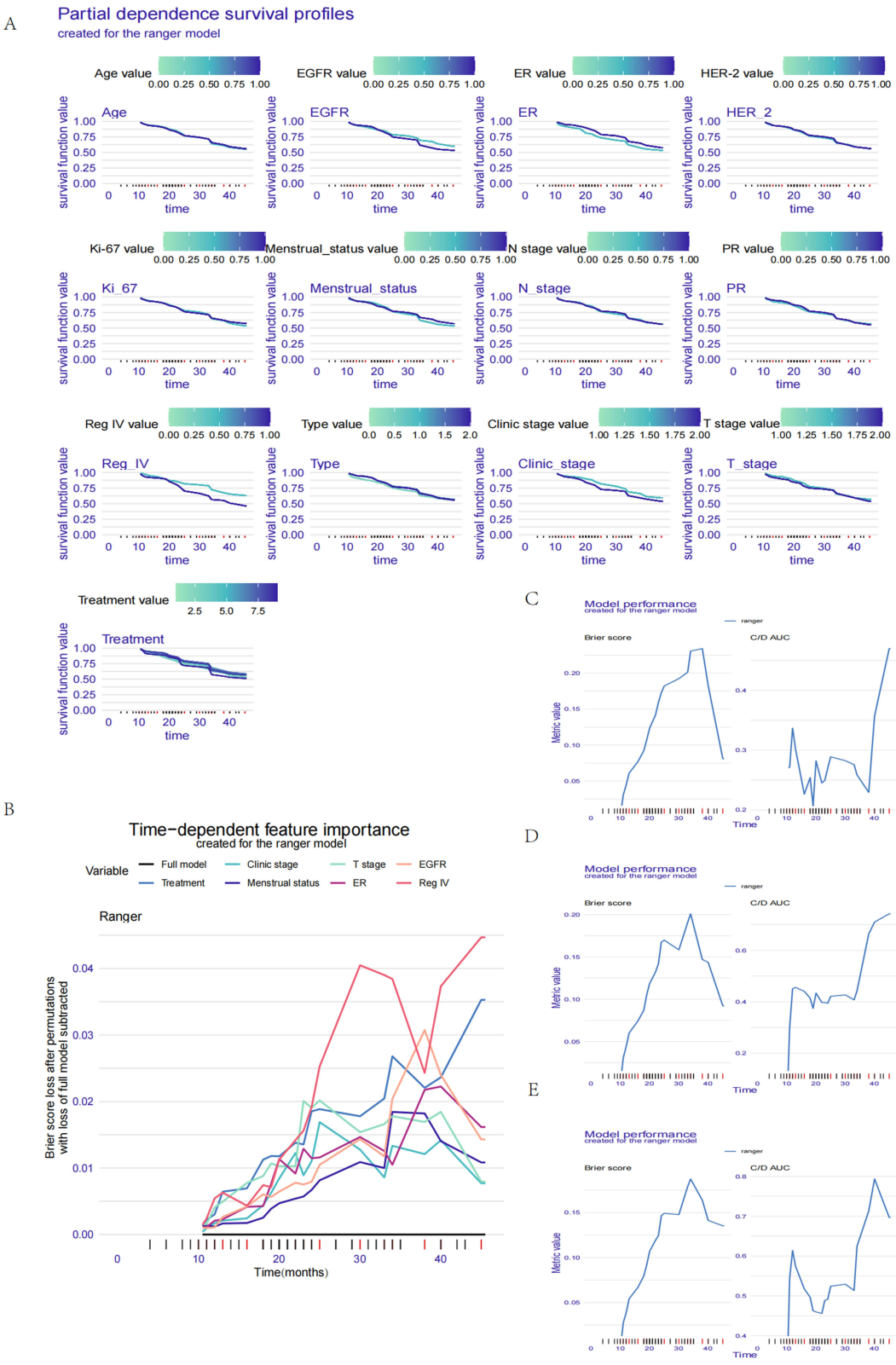

Figure 3 The interaction of EGFR with SHAP and Reg IV. (**A**). Interpretability analysis based on model variables involved depicting the survival dependency graph of each variable. (**B**). The time-dependent ROC curve depicted variable importance. The different colors in the X coordinate identify the variables that contribute the most to the whole model at that point in time, for example, Reg IV contributes the most to the model at 30 months, and ER contributes the most to the model at 40 months. (**C–E**) These figures illustrate the time-dependent ROC curve depicting variable importance for the Reg IV model, the model of Reg IV combined with EGFR, and the model of Reg IV combined with EGFR and TN, respectively.

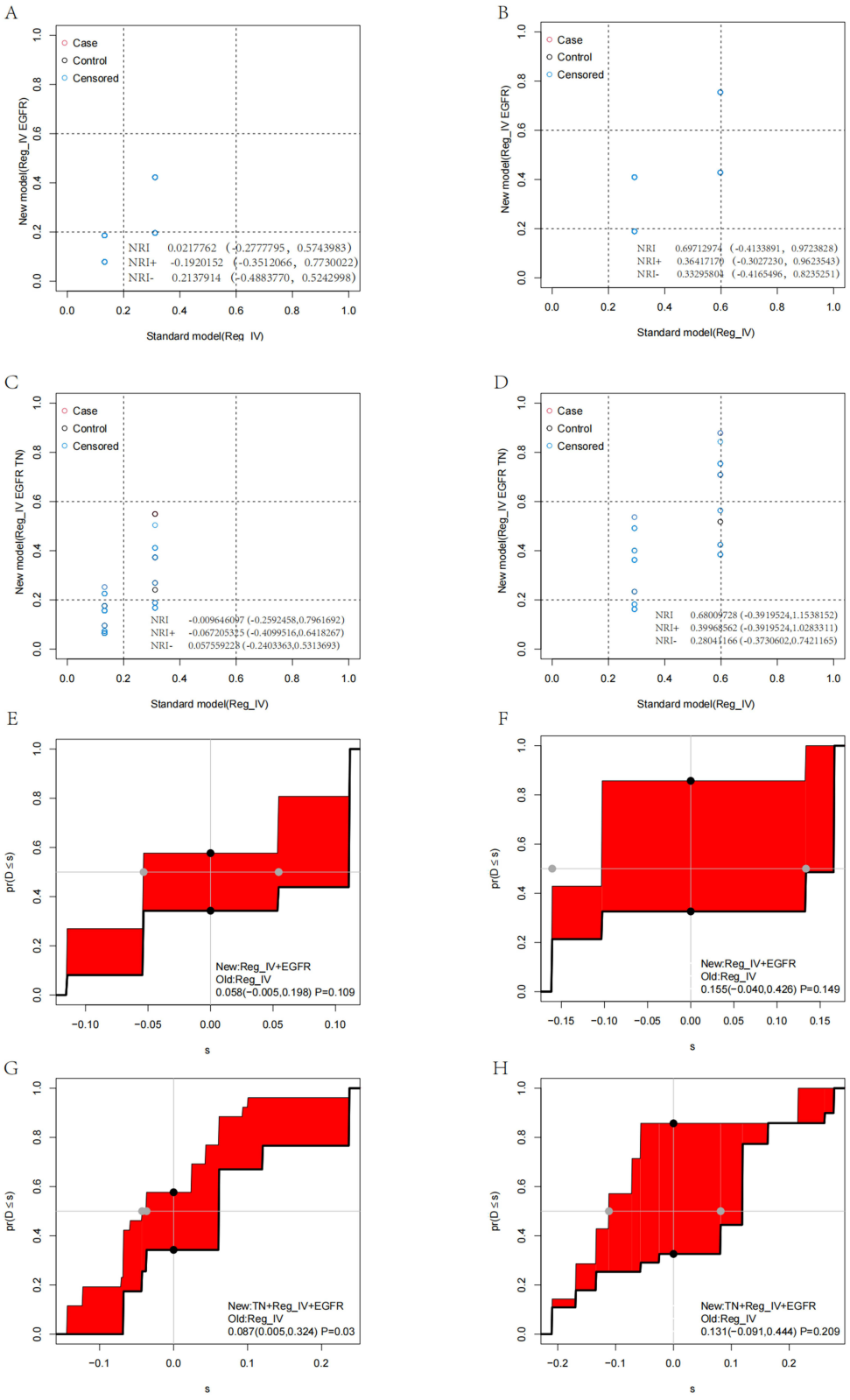

Figure 4 NRI and IDI. (**A**–**D**) represent the NRI curves at 24 and 36 months for the models of Reg IV, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging, respectively (Note: There are no event and control groups in (**A** and **B**), only the deletion group, because no such patients were present at that point in time). (**E–H)** represent the IDI curves at 24 and 36 months for the models of Reg IV, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging.

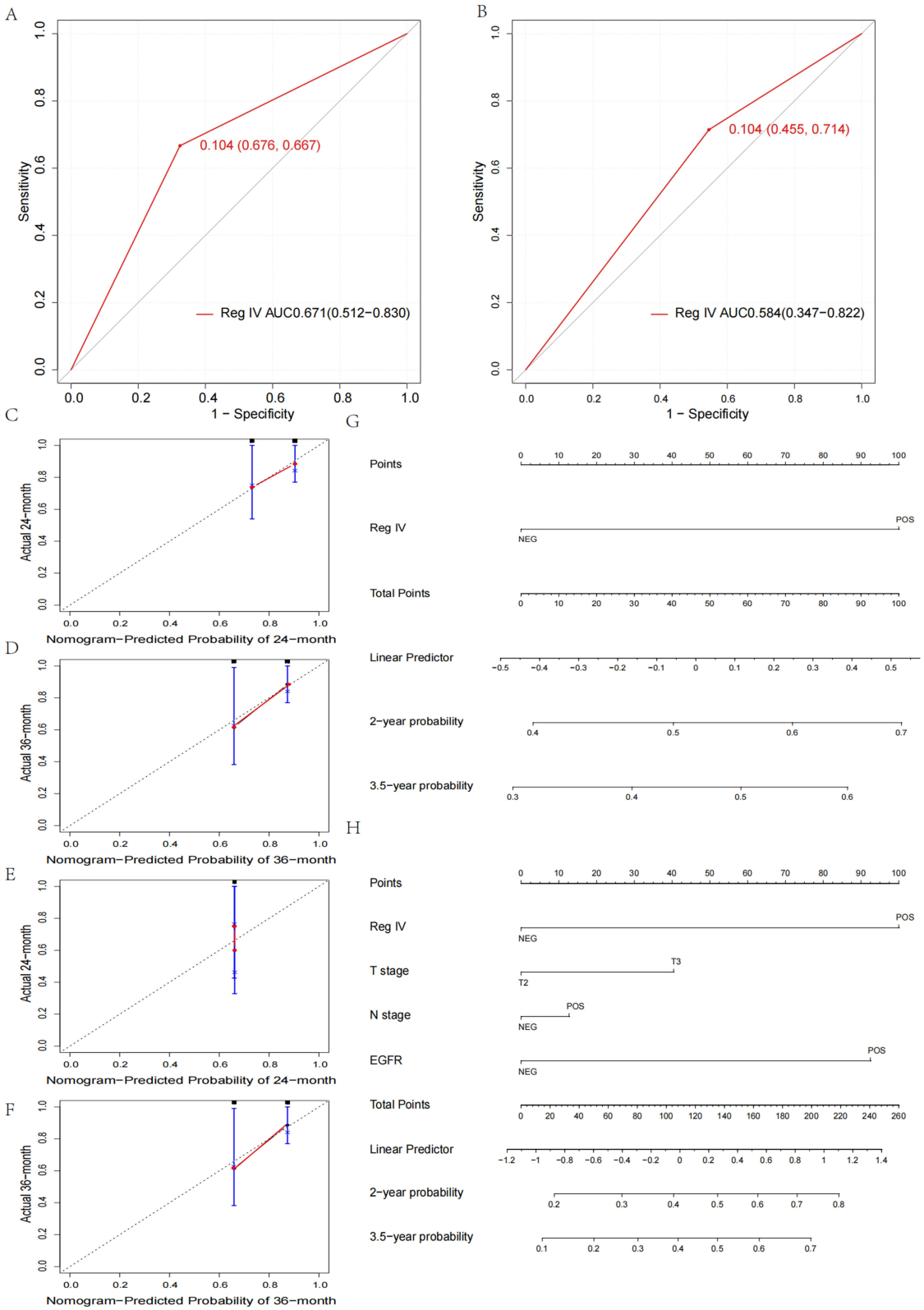

Figure 5 ROC curve and calibration curve of the training and validation sets. (**A** and **B)** represent the ROC curves of Reg IV in the training and validation sets, respectively. (**C–F**) represent the calibration curves of Reg IV for 24 and 36 months in the training and validation sets, respectively. (**G** and **H)** depict the nomogram chart of the Reg IV model and the model of Reg IV combined with EGFR and TN staging.

24 months to support the calibration curve in distinguishing between different risk levels. The efficiency of Reg IV models was assessed across different datasets (Figure 5G and H):

Training Set: Accuracy: 0.6735 (95% CI: 0.5246, 0.8005) No Information Rate: 0.5918, P-value [Acc > NIR]: 0.15441, Kappa: 0.2794, McNemar's Test P-value: 0.08012, Sensitivity: 0.4000, Specificity: 0.8621, Positive Predictive Value: 0.6667, Negative Predictive Value: 0.6757, Prevalence: 0.4082, Detection Rate: 0.1633, Detection Prevalence: 0.2449, Balanced Accuracy: 0.6310.

Verification Set: Accuracy: 0.5556 (95% CI: 0.3076, 0.7847), No Information Rate: 0.6111, P-value [Acc > NIR]: 0.7680, Kappa: 0.1529, McNemar's Test P-value: 0.2888, Sensitivity: 0.4545, Specificity: 0.7143, Positive Predictive Value: 0.7143, Negative Predictive Value: 0.4545, Prevalence: 0.6111, Detection Rate: 0.2778, Detection Prevalence: 0.3889, Balanced Accuracy: 0.5844.

Discussion

The increasing prevalence of NACT in breast cancer treatment has drawn attention to both patients achieving pCR and those classified as non-pCR. Non-pCR status is an independent prognostic indicator linked with unfavorable outcomes.[13] However, the specific high-risk factors influencing prognosis in non-pCR patients remain unclear. Reg IVR, a member of the REG family encoding the Reg IV protein, was first identified in 2001 through high-throughput sequencing of cDNA libraries from ulcerative colitis (UC) samples. Reg IV exhibits increased expression, which fosters tumor cell invasion and migration.[14,15]

This study included 67 patients eligible for statistical analysis, representing a relatively small cohort. Preliminary analysis revealed a correlation between Reg IV and the prognosis of non-pCR patients following T2-3 neoadjuvant chemotherapy, a finding that has not been previously reported. These results are reported with caution, and it is recommended to expand the dataset and conduct multicenter studies to further validate these findings and address potential biases. Within our study group, 19 patients (28.3%) experienced disease progression events, with 6 (32%) and 13 (68%) testing negative and positive for Reg IV, respectively. Among patients experiencing disease progression, the prevalence of positive Reg IV expression significantly surpassed that of negative Reg IV expression. This observation concurs with the known propensity of Reg IV to stimulate cell proliferation and impede apoptosis, consequently augmenting breast cancer malignancy and correlating with a poorer prognosis in patients who are Reg IV-positive.

The prognosis of patients with varying expression levels of Reg IV and EGFR was thoroughly analyzed in this study. Results indicated that patients positive for Reg IV exhibited a worse prognosis compared to those negative for Reg IV. Similarly, patients positive for EGFR demonstrated a poorer prognosis compared those negative for EGFR. Also, patients co-positive for both Reg IV and EGFR displayed the most unfavorable prognosis among all patient groups, while those co-negative for both exhibited the best prognosis. Patients with either positive Reg IV and negative EGFR, or negative Reg IV and positive EGFR, revealed similar prognoses. Kaplan–Meier survival analysis revealed certain characteristics; however, as this method could not effectively account for the influence of other factors, Cox regression analysis was subsequently conducted. This analysis yielded consistent results, indicating that Reg IV protein functions as an independent prognostic indicator. Despite limited research on Reg IV in breast cancer, similar patterns have been observed in other tumor types.[16,17]

In contrast to earlier investigations,[18] where the Reg IV protein was identified as a potent stimulator of the EGFR/phosphoinositide 3-kinase/Akt/AP-1 signaling pathway in human colon cancer cell lines and was linked to EGFR expression in patients with colon cancer, our dataset did not reveal a significant correlation between Reg IV and EGFR expression (correlation coefficient = 0.4, $P > 0.05$).[19] However, patients exhibiting positivity for both Reg IV and EGFR experienced the poorest prognosis, indicating a partial overlap in the roles played by Reg IV and EGFR in influencing patient outcomes.

We considered various perspectives in our analysis of the role of Reg IV in predicting the prognosis of non-pCR patients. Both univariate and multivariate Cox regression analyses highlighted Reg IV as an independent risk factor, with the EGFR value serving as a crucial threshold for distinguishing between positive and negative results. Through 10-fold cross-validation of the data, the fluctuation of the risk ratio ranged from a minimum of 1.94 (1.08, 3.47) to a maximum of 3.82 (2.82, 5.19), resulting in pooled hazard ratios of 2.92 (2.57, 3.31) and 2.78 (2.35, 3.30), respectively. The I^2 value

was 39%, indicating moderate heterogeneity. Notably, a high expression of Reg IV was consistently linked to adverse prognosis across analyses.

Relevant studies have identified residual tumor cellularity (RTC) and pCR following neoadjuvant chemotherapy as prognostic factors associated with enhanced outcomes in breast cancer. Tumor stage prior to neoadjuvant chemotherapy, breast cancer subtype, and vascular invasion are significant and independent factors related to pCR. Patients achieving pCR with RTC < 40% generally experience longer DFS, distant disease-free survival (DDFS), and OS compared to those with non-pCR.[20] This study, along with our findings, examines prognostic factors from two distinct dimensions: tumor characteristics and residual tumor cell quantity. Variations in Reg IV protein expression levels predict different prognoses, while the amount of residual tumor cells may reflect both the original tumor characteristics and the effectiveness of the treatment plan. Ultimately, the prognosis of patients is determined by the conformity and characteristics of the tumors, a perspective supported and expanded by various studies.

Our assessment of Reg IV alone, Reg IV combined with EGFR, and Reg IV combined with EGFR and TN staging revealed that incorporating Reg IV with other variables did not significantly enhance predictive power for DFS. We hypothesize that Reg IV may reflect tumor characteristics rather than tumor burden in patients, thereby failing to comprehensively depict the current disease status. This aligns with similar conclusions drawn from our previous research.[21] Consequently, we proceeded to investigate the stability and prognostic influences of Reg IV at different time points. Subsequently, we discovered that the models combining Reg IV with EGFR, and Reg IV with EGFR and TN staging, exhibited relatively stable efficacy across different time points. Notably, these combined models demonstrated a higher AUC value compared to the Reg IV model alone.

We investigated the potential interaction between Reg IV and EGFR. Specifically, we assessed the impact of EGFR on the prognosis of patients with either high or low expression levels of Reg IV. However, we discovered that EGFR had no statistically significant impact on the prognosis of patients, regardless of their Reg IV expression levels. However, an interesting observation emerged regarding the relationship between Reg IV expression and prognosis among patients with differing EGFR expression levels. In patients with high EGFR expression, Reg IV significantly impacted prognosis (β = 1.599, SE = 0.792, Wald = 2.02, P = 0.043). Conversely, patients with low EGFR expression did not exhibit such a statistical difference (β = 0.648, SE = 0.731, Wald = 0.89, P = 0.375). Further examination of the interaction between EGFR and Reg IV revealed no significant interaction between the two (β = 0.575, SE = 1.0, Wald = 0.58, P = 0.565).

Some patients exhibited a low EGFR expression but a high Reg IV expression. However, this observation did not demonstrate any statistical significance in terms of prognosis. We speculate two potential explanations for this finding. Firstly, it could be attributed to the relatively small sample size in this study, which may not fully represent the broader population and could introduce bias into the analysis. Secondly, besides the EGFR pathway, Reg IV may be involved in a signaling pathway independent of EGFR. However, this pathway may not fully exert its function when EGFR expression is low. In essence, there is a possibility that Reg IV and EGFR operate independently of each other. Nonetheless, when both are present, they might synergistically contribute to the increased malignancy of the tumor. This hypothesis warrants further validation through additional research and a larger sample size.

This analysis, which included variable importance assessment through SHAP and modeling across different datasets, provides a comprehensive explanation of the model. SHAP analysis elucidates the degree to which each variable contributes to the predictions of the model, thereby clarifying the importance of each variable in determining the prediction outcomes. This approach helps in identifying and selecting significant variables. The findings consistently confirm that Reg IV serves as an independent risk factor for the prognosis of patients with T2-3 breast cancer, especially those who exhibit non-pCR following NACT. Additionally, Reg IV serves as a reliable predictor in this context. Although enhancements in the AUC value are noted with increases in T and N stages, the advancement of the overall model lacks statistical significance, with only one instance of significance detected at the 24-month IDI (0.087, 95% CI: 0.005 to 0.324, P = 0.03). Notably, no disparities were observed in the NRI. We attribute these variances in performance metrics to potential errors stemming from the limited sample size. However, acknowledging the potential for enhanced predictive accuracy through the incorporation of easily accessible clinical data, we advocate for the integration of such variables to bolster the prognostic capabilities of the model.

Despite the insights gleaned from this study, several limitations must be acknowledged. Firstly, our dataset is confined to a specific temporal and institutional context, curtailing its generalizability to broader populations. For example, when

examining the stability of the model in the 24-month verification set, the risk stratification was compromised due to the limited number of outcome events. This limitation underscores the need for further research with an expanded sample size and multi-center studies to enhance the robustness of the findings. Secondly, the relatively small sample size may introduce bias and compromise result accuracy. Finally, clinical decision-making entails multifaceted considerations, emphasizing the need for a comprehensive assessment beyond endpoint events alone.

Moving forward, we aim to further elucidate the role of Reg IV in patients with breast cancer, with the goal of detecting prospective predictive markers and therapeutic targets. This entails broadening our scope to encompass diverse patient groups and assessing additional clinical parameters to enhance prognostic assessment and treatment stratification.

Conclusion

In conclusion, the results of this study indicate that Reg IV serves as an independent risk factor and predictor for unfavorable prognosis among patients with T2-3 breast cancer who do not achieve pCR following NACT. These findings underscore the potential clinical use of Reg IV as a prognostic marker in guiding treatment decisions. However, further research is needed to validate its efficacy and incorporate it into routine clinical practice.

Abbreviations

Reg IV, Regenerating islet-derived family member IV; pCR, Pathologic Complete Response; TNM, Tumor-Node-Metastasis; FISH, Fluorescence In Situ Hybridization; UICC, Union for International Cancer Control; AJCC, American Joint Committee on Cancer; NACT, Neoadjuvant Chemotherapy; AUC, Area Under Curve; NRI, Net Reclassification Improvement; IDI, Integrated Discrimination Improvement; DCA, Decision Curve Analysis; CIC, Clinical impact curve; Xgboost, eXtreme Gradient Boosting; ROC, Receiver Operating Characteristic Curve; SHAP, SHapley Additive exPlanation.

Data Sharing Statement

The datasets generated and analysed during the current study are not publicly available but are available from the corresponding author (Jun Shen) on reasonable request.

Ethics Approval and Consent to Participate

This study was conducted in accordance with the declaration of Helsinki. This study was conducted with approval from the Ethics Committee of the First People's Hospital of Lianyungang (KY-20230410002-01). A written informed consent was obtained from all participants.

Consent for Publication

Consent for publication was obtained from every individual whose data are included in this manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Zheng HC, Xue H, Zhang CY. REG4 promotes the proliferation and anti-apoptosis of cancer. *Front Cell Dev Biol*. 2022;10:1012193. doi:10.3389/fcell.2022.1012193
2. Zhang J, Zhu Z, Miao Z, et al. The clinical significance and mechanisms of reg4 in human cancers. *Front Oncol*. 2021;10:559230. PMID: 33489872; PMCID: PMC7819868. doi:10.3389/fonc.2020.559230
3. Mougalian SS, Hernandez M, Lei X, et al. Ten-year outcomes of patients with breast cancer with cytologically confirmed axillary lymph node metastases and pathologic complete response after primary systemic chemotherapy. *JAMA Oncol*. 2016;2(4):508–516. doi:10.1001/jamaoncol.2015.4935
4. Abdel-Razeq H, Khalil H, Assi HI, Dargham TB. Treatment strategies for residual disease following neoadjuvant chemotherapy in patients with early-stage breast cancer. *Curr Oncol*. 2022;29(8):5810–5822. doi:10.3390/curroncol29080458
5. Mitani Y, Oue N, Matsumura S, et al. Reg IV is a serum biomarker for gastric cancer patients and predicts response to 5-fluorouracil-based chemotherapy. *Oncogene*. 2007;26(30):4383–4393. doi:10.1038/sj.onc.1210215

6. Nanakin A, Fukui H, Fujii S, et al. Expression of the REG IV gene in ulcerative colitis. *Lab Invest*. 2007;87(3):304–314. doi:10.1038/labinvest.3700507
7. Bishnupuri KS, Luo Q, Murmu N, Houchen CW, Anant S, Dieckgraefe BK. Reg IV activates the epidermal growth factor receptor/Akt/AP-1 signaling pathway in colon adenocarcinomas. *Gastroenterol*. 2006;130(1):137–149. doi:10.1053/j.gastro.2005.10.001
8. Shuai Y, Ma L. Prognostic value of pathologic complete response and the alteration of breast cancer immunohistochemical biomarkers after neoadjuvant chemotherapy. *Pathol Res Pract*. 2019;215(1):29–33. doi:10.1016/j.prp.2018.11.003
9. Singletary SE, Allred C, Ashley P, et al. Staging system for breast cancer: revisions for the 6th edition of the AJCC Cancer Staging Manual. *Surg Clin North Am*. 2003;83(4):803–819. doi:10.1016/S0039-6109(03)00034-3 PMID: 12875597.
10. Wolff AC, Hammond ME, Hicks DG, et al. American society of clinical oncology; college of American pathologists. recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline update. *J Clin Oncol*. 2013;31(31):3997–4013. doi:10.1200/JCO.2013.50.9984 Epub 2013 Oct 7. PMID: 24101045.
11. Pepe MS, Fan J, Feng Z, Gerds T, Hilden J. The net reclassification index (NRI): a misleading measure of prediction improvement even with independent test data sets. *Stat Biosci*. 2015;7(2):282–295. doi:10.1007/s12561-014-9118-0
12. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *JAMA*. 2017;318(14):1377–1384. doi:10.1001/jama.2017.12126 PMID: 29049590.
13. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014;384(9938):164–172. doi:10.1016/S0140-6736(13)62422-8 PMID: 24529560.
14. Hartupee JC, Zhang H, Bonaldo MF, Soares MB, Dieckgraefe BK. Isolation and characterization of a cDNA encoding a novel member of the human regenerating protein family: reg IV. *Biochim Biophys Acta*. 2001;1518(3):287–293. doi:10.1016/S0167-4781(00)00284-0
15. Kämäräinen M, Heiskala K, Knuutila S, Heiskala M, Winqvist O, Andersson LC. RELP, a novel human REG-like protein with up-regulated expression in inflammatory and metaplastic gastrointestinal mucosa. *Am J Pathol*. 2003;163(1):11–20. doi:10.1016/S0002-9440(10)63625-5
16. Ohara S, Oue N, Matsubara A, et al. Reg IV is an independent prognostic factor for relapse in patients with clinically localized prostate cancer. *Cancer Sci*. 2008;99(8):1570–1577. doi:10.1111/j.1349-7006.2008.00846.x PMID: 18754868.
17. Hu Y, Pan C, Hu J, Zhang S. The role of Reg IV in colorectal cancer, as a potential therapeutic target. *Contemp Oncol*. 2015;19(4):261–264.
18. Wolff RA, Evans DB, Gravel DM, et al. Phase I trial of gemcitabine combined with radiation for the treatment of locally advanced pancreatic adenocarcinoma. *Clin Cancer Res*. 2001;7(8):2246–2253. PMID: 11489798.
19. Li XH, Zheng Y, Zheng HC, et al. REG IV overexpression in an early stage of colorectal carcinogenesis: an immunohistochemical study. *Histol Histopathol*. 2010;25(4):473–484. doi:10.14670/HH-25.473 PMID: 20183800.
20. Gentile D, Sagona A, De Carlo C, et al. Pathologic response and residual tumor cellularity after neo-adjuvant chemotherapy predict prognosis in breast cancer patients. *Breast*. 2023;69:323–329. doi:10.1016/j.breast.2023.03.016
21. Shen J, Wang M, Li F, Yan H, Wang R, Zhou J. Establishment and validation of a model for disease-free survival rate prediction using the combination of microRNA-381 and clinical indicators in patients with breast cancer. *Breast Cancer*. 2022;14:375–389. doi:10.2147/BCTT.S383121