RESEARCH

Prognostic significance of the novel immunonutritional marker of cholesterol-to-lymphocyte ratio in patients with non-metastatic breast cancer

Li-Wen Wan[1†], Chao Zhang[2†], Yu-Ling Zhang[3†], Fei Lin[4*†], Xin Hua[5*†] and Wen Xia[6*†]

Abstract

Background Although there is a strong correlation between the novel cholesterol-to-lymphocyte ratio (CLR) and tumor survival, its prognostic significance in breast cancer (BC) is unknown. After analyzing the relationship between CLR and the overall survival (OS) of patients with BC, we created a predictive model.

Methods Following retrospective enrollment, 1316 patients with BC were randomized into two cohorts: validation ($n = 392$) and training ($n = 924$). Distinct factors within the training dataset were identified for OS by univariate and multivariate Cox analyses; two-tailed P-value < 0.05 were considered to indicate statistical significance. On this premise, we developed novel signals for survival prediction and utilized the calibration curve, receiver operating characteristic curves, and concordance index (C-index) to validate their efficacy across both datasets.

Results Patients with BC were categorized into two categories with differing prognoses based on the CLR score [hazard ratio = 0.492; 95% confidence interval (CI): 0.286–0.846, $P = 0.009$]. A prediction nomogram was created based on multivariate analysis, which showed that N stage, postoperative pathological categorization, and CLR score were all independently correlated with OS. In the training [C-index = 0.831 (95% CI: 0.788–0.874)] and validation [C-index = 0.775 (95% CI: 0.694–0.856)] cohorts, the nomogram demonstrated favorable performance in predicting OS. In both the training and validation cohorts, it outperformed the traditional staging system [C-index = 0.702 (95% CI: 0.623–0.782)] and [C-index = 0.709 (95% CI: 0.570–0.847)]. The accurate prediction by the signature was further demonstrated by the time-dependent receiver operating characteristic curves.

†Li-Wen Wan, Chao Zhang and Yu-Ling Zhang contributed equally to this work and share first authorship.

†Fei Lin, Xin Hua and Wen Xia contributed equally to this work and share senior authorship.

*Correspondence:
Fei Lin
linfei06050@gmail.com
Xin Hua
hx12914@rjh.com.cn
Wen Xia
xiawen@sysucc.org.cn

Full list of author information is available at the end of the article

Conclusions The novel immunonutritional marker CLR could function as a simplified, cost-effective, easily accessible, non-invasive, and readily promotive prognostic indicator for patients with early-stage BC and demonstrates superior predictive power than the traditional staging system.

Keywords Cholesterol-to-lymphocyte ratio, Breast cancer, Immunonutritional marker, Prognosis, Nomogram

Introduction

Breast cancer (BC) stands as a prevalent malignant tumor disease in women [1, 2]. Despite continuous enhancements in treatment methods, BC remains the primary cause of cancer-related mortality among female patients globally [3, 4]. In clinical settings, clinicians typically make therapy decisions based on the molecular type of BC, taking into account markers such as Ki67, human epidermal growth factor receptor 2 (HER2), progesterone receptor (PR), and estrogen receptor (ER). However, patient heterogeneity may result in significantly varied outcomes for patients with BC at the same stage and undergoing the same treatment [5]. Thus, identifying alternative biomarkers to enhance the diagnostic system and developing more practical predictive models for personalized therapeutic guidance becomes imperative [6–9].

The association between cancer progression and alterations in lipid metabolism and inflammation has been shown [10]. Cancer cells augment their uptake and accumulation of cholesterol, phospholipids, and fatty acids as a means of sustaining life in a nutrient-deficient microenvironment. Zhou et al. [11]. discovered in a recent study that the cholesterol-to-lymphocyte ratio (CLR), a lipid metabolism and inflammatory state indicator, is connected with the long-term prognosis of colorectal cancer. Moreover, it exhibited greater sensitivity and specificity than the neutrophil-lymphocyte ratio, a widely used inflammatory marker [10]. Deng et al. discovered that the unfavourable outcomes in colorectal cancer liver metastasis patients who are receiving simultaneous resection of the primary lesion and liver metastases can be predicted by the preoperative CLR level [12]. Matsubara discovered that the cholesterol-lymphocyte score is closed related to the prognosis of patients with gastric cancer [13]. Several nutritional-inflammatory prognostic indexes have been associated with survival outcomes in the context of BC, including the platelet-lymphocyte ratio, systemic inflammation score, neutrophil-lymphocyte ratio, monocyte-lymphocyte ratio, Controlling Nutritional Status, and systemic immune-inflammation index. These indices have been created and therapeutically utilized for a variety of malignancies, including BC [14–17]. However, the predictive capability of the CLR score for survival outcomes in patients with BC remains unclear.

This study aimed to investigate the prognostic significance of the CLR score and develop a model that can provide individualized survival predictions for patients with early-stage BC.

Methods

Patients

Between December 2010 and October 2012, a total of 1316 patients who were first diagnosed with BC at the Sun Yat-sen University Cancer Center (SYSUCC; Guangzhou, China) were randomly divided into two groups: a verification dataset and a training dataset. The allocation was in a 7:3 ratio. The following were the criteria for inclusion: (1) age≥18 years old; (2) a histological diagnosis of invasive BC; (3) the absence of distant metastases (to the lungs, bones, liver, or brain); (4) accessible baseline laboratory data and particular follow-up data. The criteria for exclusion were the following: (1) Pregnant or breastfeeding patients; (2) Patients diagnosed with ductal carcinoma in situ; (3) Patients with synchronous malignancies; (4) Patients who, in the last three months, have taken any medicine inducing an inflammatory or immunological response; and (5) Patients with any inflammatory disease, including autoimmune diseases.

This study was approved by the Research Ethics Committee of SYSUCC, and all procedures involving human participants adhered to the organization's ethical principles, the Helsinki Declaration (1604), its subsequent amendments, and analogous ethical standards. The patients' need to provide written informed consent was not enforced in this retrospective study.

Collection of data and categorization of variables

Manual extraction of all clinicopathological characteristics of the patients was performed utilizing the SYSUCC electronic medical records system. The following characteristics were assessed: the age of the patient at the time of diagnosis, pathological grade, T and N stage, ER, PR, and HER2 status, lymphatic vessel invasion, Ki67 index (which was calculated using an MIB1 monoclonal antibody from ZSGBBIO, Beijing, China by two pathologists), and postoperative pathological classification. Pretreatment examinations were performed on all participants, comprising a physical examination, hematological and biochemical testing, and an inquiry into their medical history. From the electronic medical records system, blood parameters, particularly cholesterol and lymphocyte levels within one week before any antitumor therapy, were collected.

Wan *et al. BMC Cancer*   (2024) 24:914

Page 3 of 10

BC specimens were classified as ER-positive or PR-positive, correspondingly, if immunohistochemical analysis revealed that over 1% of the tumor nuclei were positive for PR or ER. Concerning *HER2* status, only tumor cells that obtained a 3+immunohistochemistry score or a 2+fluorescence in situ hybridization score for the *HER2* gene were classified as *HER2*-positive [15]. CLR was calculated by dividing total serum cholesterol by the lymphocyte count. This continuous laboratory variable CLR was then categorized, and the optimal cutoff value (3.93) was determined using maximally selected rank statistics analysis [16].

Outcome and follow-up

Overall survival (OS) was defined as the period from the time of diagnosis and the patient's death from any cause or the last follow-up. The disease-free survival time was defined as the time between the patient's diagnosis and the commencement of the first illness progression or fatality from any cause. The SYSUCC electronic medical record system was used to manually collect the patients' clinicopathological data. Telephone interviews were used to track the patients' various medical conditions.

Statistical analysis

Continuous variables were presented as median values with interquartile ranges, and categorical variables were listed as frequencies with percentages. Fisher's exact or chi-square tests were used to compare variables between the two cohorts. The optimal CLR cutoff value was determined using the "maxstat" software using maximally chosen rank statistics with survival status as the endpoint. By utilizing the log-rank tests, the survival curves produced by the Kaplan–Meier technique for the validation cohort and the training cohort were compared. To examine the proportional hazards hypothesis, Schoenfeld residuals were utilized. The Cox proportional hazards model was employed to develop and validate the prognostic model through univariate and multivariate analyses. In the multivariate analysis, variables that met the criterion of having a P-value of <0.05 in the univariate analysis were included to identify independent risk factors. A nomogram was constructed and utilized to represent a prognostic model constructed from the independent risk variables revealed in the multivariate analysis of the training cohort (A nomogram consists of a set of n scales, n-1 of which are graduated linearly. The nth scale can be linear or nonlinear. Each scale represents one of the variables involved in the equation. To read the value of an unknown variable, a line is drawn between the known values on the fixed scales. The position of the line where it intersects the unknown scale indicates the value of the unknown variable.). To assess the discriminative capability of the model, the "rms" package was employed

to calculate Harrell's concordance index (C-index), "timeROC" for time-dependent receiver operative characteristics (tROC) (tROC is a statistical tool used primarily in survival analysis to evaluate the accuracy of a predictive model at specific points in time. It takes into account the time-varying nature of the outcomes, such as disease progression or death, in longitudinal data.), and "ggDCA" for decision curve analysis (DCA) in the training and validation cohorts (DCA is a method used in evidence-based medicine to evaluate the clinical utility of diagnostic tests, prognostic models, and treatment strategies.DCA incorporates the clinical consequences of decisions made based on predictions. It quantifies the net benefit of a model or strategy by comparing it to the extremes of 'treat all' and 'treat none', taking into account the threshold probabilities at which patients or clinicians would opt for treatment). The area under the curve (AUC) of the time-dependent ROC analysis, the C-index, the DCA curve, and the calibration curve were utilized to evaluate the performance of the nomogram (A calibration curve is a graphical tool used in statistics to assess the reliability of predictive models, particularly those used for probabilistic forecasting. It compares the actual outcomes against the predicted probabilities from a model.). R 4.2.1 was utilized to conduct the statistical analysis.

Results

Patients' characteristic

We enrolled 1316 patients with early BC, who were randomized in a 7:3 ratio into training and validation datasets (924 vs. 392, respectively). Table 1 shows their clinicopathological characteristics, which were comparable between the two groups. In the training cohort, 372 (40.3%) patients, and in the validation cohort, 150 (38.3%) patients were aged over 50 years. The majority of patients ($n=1114$) were pathologically diagnosed with invasive ductal carcinoma. T2 (54.1%) and N0 (51.6%) patients made up roughly half of the patients, with *HER2*-negative patients accounting for 88.6% (150 patients). The majority of patients ($n=1114$) were pathologically diagnosed with invasive ductal carcinoma.

In the whole cohort of 1316 patients, 124 death events were observed, with 88 in the training cohort and 36 in the validation cohort. The 1-, 3-, and 5-year OS rates for the entire cohort were 97.18%, 85.60%, and 70.52%; the training cohort's rates were 97.50%, 85.45%, and 69.92%; and the validation cohort's rates were 96.43%, 85.97%, and 71.94%. An analysis of the training and validation cohorts revealed no statistically significant difference in OS (Fig. 1, $P=0.790$).

Wan *et al. BMC Cancer* (2024) 24:914

Page 4 of 10

Table 1 Patient demographics and clinical characteristics between the training and validation cohorts

Characteristic	Train set	Validation set	*P*-value
Age			0.539
≥ 50 years	372 (40.3%)	150 (38.3%)	
< 50 years	552 (59.7%)	242 (61.7%)	
Menstrual status			0.426
Post	396 (42.9%)	158 (40.3%)	
Pre	528 (57.1%)	234 (59.7%)	
Histological type			0.200
IDC	774 (83.8%)	340 (86.7%)	
Other	150 (16.2%)	52 (13.3%)	
Subtype			0.266
HER2	106 (3.06%)	44 (11.2%)	
LA	263 (64.9%)	98 (25.0%)	
LB1	274 (32.1%)	138 (35.2%)	
LB2	171 (18.5%)	62 (15.8%)	
TNBC	110 (11.9%)	50 (12.8%)	
Endocrine therapy			0.571
No	239 (25.9%)	108 (27.6%)	
Yes	685 (74.1%)	284 (72.4%)	
Adjuvant chemotherapy			0.348
No	180 (19.5%)	67 (17.1%)	
Yes	744 (80.5%)	325 (82.9%)	
Radiotherapy			0.759
No	681 (73.7%)	285 (72.7%)	
Yes	243 (26.3%)	107 (27.3%)	
Target therapy			0.962
No	858 (92.9%)	363 (92.6%)	
Yes	66 (7.14%)	29 (7.40%)	
T stage			0.900
T1	334 (36.1%)	134 (34.2%)	
T2	500 (54.1%)	221 (56.4%)	
T3	46 (4.98%)	19 (4.85%)	
T4	44 (4.76%)	18 (4.59%)	
N stage			0.177
N0	477 (51.6%)	211 (53.8%)	
N1	256 (27.7%)	91 (23.2%)	
N2	116 (12.6%)	47 (12.0%)	
N3	75 (8.12%)	43 (11.0%)	
CLR			0.334
≥ 3.93	826 (89.4%)	358 (91.3%)	
< 3.93	98 (10.6%)	34 (8.67%)	

Abbreviations IDC = invasive ductal carcinoma; HER2 = human epidermal growth factor receptor 2; LA = luminal A; LB1 = luminal B1; LB2 = luminal B2; TNBC = triple-negative breast cancer; CLR = cholesterol-to-lymphocyte ratio

CLR score prognostic value for OS in BC

In both the training and validation cohorts, Kaplan–Meier curves revealed that patients in the low-CLR group had a significantly higher survival rate than those in the high-CLR group [Fig. 2A, hazard ratio = 0.492; 95% CI: 0.286–0.846, *P* = 0.009; Fig. 2B, hazard ratio = 0.352; 95% CI: 0.154–0.803, *P* = 0.009]. As shown in the table, there is no significant difference of CLR among different subtypes (Supplementary Table 1, *P* = 0.562). As shown

in the figure, the Kaplan-Meier curves for OS according to CLR strata within each breast cancer subtype demonstrate that CLR exhibit statistically significant differences in certain subtypes such as HER2-enriched (*P* = 0.00088) and luminal B1 (*P* = 0.0038) (Supplementary Fig. 1).

Univariate and multivariate Cox regression analyses of OS in BC

We included age, histological type, menstruation status, T stage, N stage, HER-2 status, radiation, endocrine treatment, targeted therapy, adjuvant chemotherapy, and the CLR score in the univariate Cox regression analysis. The univariate Cox model incorporated variables, including menstrual status, T stage, N stage, histopathological type, radiotherapy, endocrine therapy, HER-2 status, and CLR score, which met the prespecified significance threshold (*P* < 0.05). These variables were then included in the multivariate Cox regression model. The multivariate Cox regression analysis revealed that OS is independently linked to histological type, N stage, and CLR score since only these variables satisfied the predetermined significance criterion (*P* < 0.05). Table 2 displays the outcomes of the univariate and multivariate Cox regression analyses.

CLR-based development of a new prognostic model

A unique nomogram-based prognostic model (Fig. 3) was constructed to forecast the 1-, 3-, and 5-year OS of an individual using the aforementioned three independent indicators derived from the multivariate modeling. The 1-year, 3-year, and 5-year OS probability could be predicted by plotting the patient's total score on the survival rate scale, which was calculated by summing the scores from each prognostic factor subtype. For instance, a patient presenting with N2 stage, a histopathological type of IDC, and a CLR score of > 3.93 would have a total point of 13.08 (3.33 + 7.88 + 1.87), corresponding to 1-, 3-, and 5-year OS probabilities of 96%, 79%, and 67%, respectively.

Evaluation of the prediction efficacy of the prognostic model

In the training and validation cohorts, the generated prognostic model exhibited excellent discriminative ability with C-indices of 0.831 (95% CI: 0.788–0.874) and 0.774 (95% CI: 0.694–0.856), respectively, outperforming the TNM staging system's C-indices of 0.703 (95% CI: 0.623–0.782) and 0.709 (95% CI: 0.571–0.847). The calibration plots for the 1-, 3-, and 5-year OS in the training (Fig. 4A) and validation cohorts revealed a high degree of concordance between the predicted and actually observed OS (Fig. 4B). Using time-dependent ROC curves to assess prognostic accuracy, it was shown that the prognostic model performed better than the standard
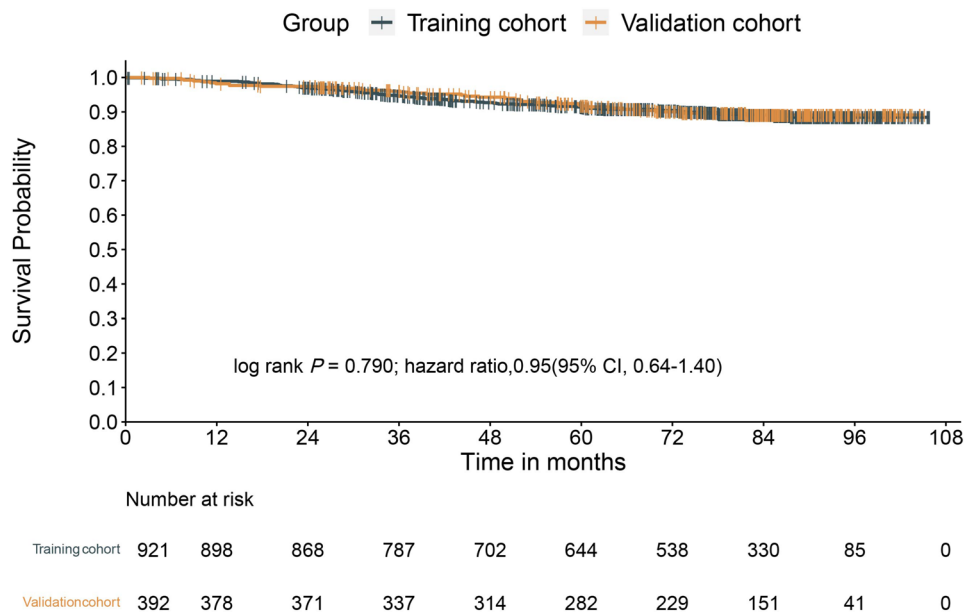
Fig. 1 Kaplan–Meier survival curves for training and validation cohorts (with unadjusted HRs). *Abbreviations* CI = confidence interval; CLR = cholesterol-to-lymphocyte ratio; HR = hazard ratios
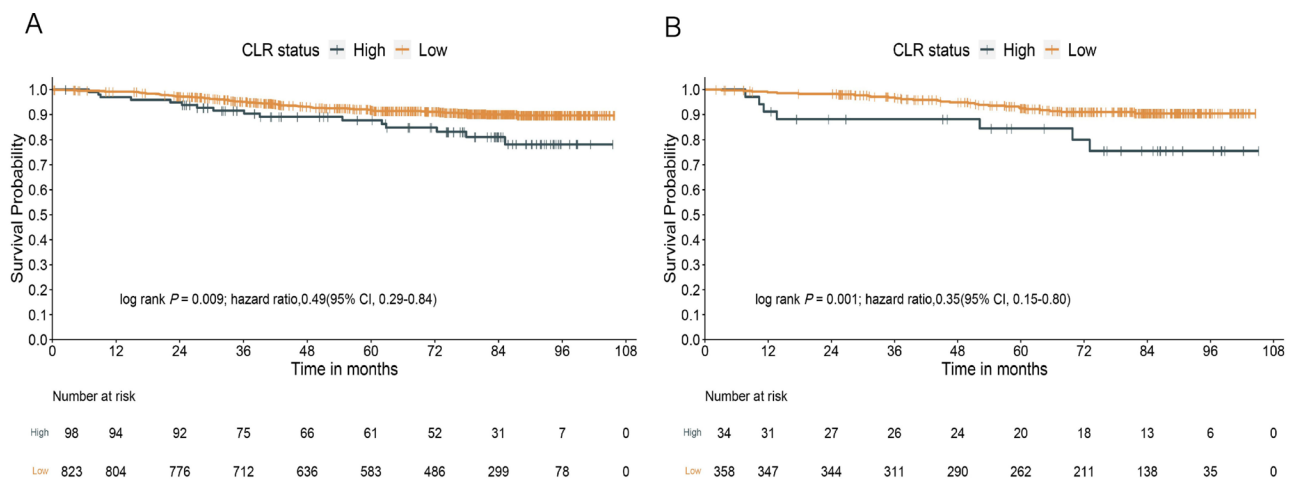

Fig. 2 Evaluation of the prognostic relevance of CLR index. Kaplan–Meier survival curves for CLR groups (with unadjusted HRs). (**A**) Training cohort survival curves. (**B**) Validation cohort survival curves. *Abbreviations* CI = confidence interval; CLR = cholesterol-to-lymphocyte ratio; HR = hazard ratios

TNM stage in both the training and validation cohorts (Fig. 4C, D). DCA curves also revealed that the new prognostic model exhibited superiority over the TNM stage in both the training and validation cohorts (Fig. 4E and F).

Discussion

An increase in the incidence of early-stage BC detection has been facilitated by advancements in medical technology and public health awareness, leading to a reduction in mortality rates [3]. Furthermore, research indicates that malnutrition and cachexia cause 20–30% of cancer-related fatalities rather than cancer itself [18, 19]. The immune system and treatment efficacy are adversely affected by malnutrition, which accelerates the

progression of diseases, localized recurrences, and distant metastases [20–22]. Therefore, the combination of nutrition support treatment and anticancer therapy is deemed essential.

TNM staging by the American Joint Committee on Cancer is the standard by which prognosis is predicted and BC therapy is directed [23]. Higher stages in patients typically imply a worse prognosis. However, previous versions of the TNM staging system had limitations as they predominantly relied on anatomical features, metastatic lymph node numbers, and distant metastases existence. These criteria fail to consider the biological diversity among patients as well as additional risk factors [23, 24], potentially impacting the prediction accuracy

Table 2 Univariate and multivariate Cox regression analyses of overall survival in the training cohort

Characteristic	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% confidence interval)	*P-value*	Hazard ratio (95% confidence interval)	*P-value*
Age				
≥ 50 years	1			
< 50 years	0.675 (0.444–1.025)	0.065		
Menstrual status				
Post	1		1	
Pre	0.649 (0.427–0.986)	0.043	0.720 (0.282–1.841)	0.493
Histological type				
IDC	1		1	
Other	0.373 (0.163–0.855)	0.020	0.405 (0.173–0.949)	0.038
Subtype				
HER2	1		1	
LA	0.172 (0.080–0.367)	< 0.001	0.419 (0.083–2.107)	0.291
LB1	0.463 (0.261–0.823)	0.009	0.799 (0.179–3.562)	0.768
LB2	0.486 (0.257–0.919)	0.026	0.757 (0.165–3.462)	0.719
TNBC	0.531 (0.260–1.087)	0.084	0.823 (0.391–1.731)	0.608
Endocrine therapy				
No	1		1	
Yes	0.492 (0.320–0.755)	0.001	0.702 (0.168–2.932)	0.628
Adjuvant chemotherapy				
No	1			
Yes	1.446 (0.786–2.658)	0.236		
Radiotherapy				
No	1		1	
Yes	1.836 (1.200–2.808)	0.005	0.760 (0.478–1.210)	0.248
Target therapy				
No	1			
Yes	1.005 (0.464–2.176)	0.99		
T stage				
T1	1		1	
T2	2.357 (1.349–4.118)	0.002	1.290 (0.719–2.316)	0.393
T3	4.327 (1.912–9.791)	< 0.001	1.515 (0.637–3.605)	0.347
T4	4.867 (2.151–11.016)	< 0.001	1.100 (0.453–2.672)	0.833
N stage				
N0	1		1	
N1	3.531 (1.748–7.135)	< 0.001	3.155 (1.546–6.436)	0.002
N2	10.157 (5.165–19.975)	< 0.001	9.334 (4.449–19.585)	< 0.001
N3	17.213 (8.682–34.128)	< 0.001	15.540 (7.279–33.171)	< 0.001
CLR				
≥ 3.93	1		1	
< 3.93	0.492 (0.286–0.846)	0.010	0.550 (0.316–0.955)	0.034

Abbreviations IDC=invasive ductal carcinoma; HER2=human epidermal growth factor receptor 2; LA=luminal A; LB1=luminal B1; LB2=luminal B2; TNBC=triple-negative breast cancer; CLR=cholesterol-to-lymphocyte ratio

of the conventional system [25, 26]. The American Joint Committee on Cancer unveiled the 8th Edition of the BC staging system in November 2016, which includes a multigene staging approach to determine the risk of recurrence, anatomical TNM staging, tumor histological grade (G), and the expression status of biomarkers (ER, PR, and HER-2). However, this approach still falls short in addressing individual heterogeneity. CLR has previously demonstrated independent prognostic significance in various cancer types, including colorectal cancer [11, 12]. Different studies employing distinct methods have shown that CLR is independently related to progression-free survival and OS, with lower CLR associated with significantly prolonged survival [11, 12]. These studies underscore the role of nutrition and immune markers in tumor prediction. However, research on CLR in BC is limited, and we hope this study helps fill this gap in the field. The CLR score and the novel prognostic model
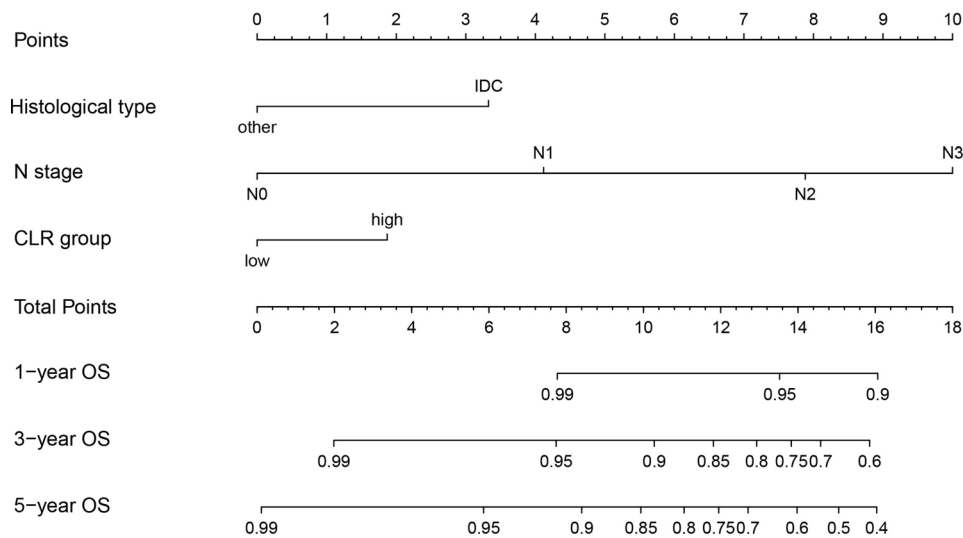
Fig. 3 Nomogram for predicting overall survival (OS) rate at 1, 3 and 5 years in patients with breast cancer (BC). The variable score can be calculated by drawing a vertical line linking the value of each parameter with the point at the top of this nomogram. Next, all scores are summed to obtain the total points score, which is plotted along the total points line, based on which the corresponding OS rates at 1, 3 years and 5 years were obtained. *Abbreviations* CLR = cholesterol-to-lymphocyte ratio; IDC = invasive ductal carcinoma; OS = overall survival

introduced in this study aim to enhance the accuracy of predicting the OS of patients with early BC. In predicting OS, time-dependent ROC curves reveal that the CLR-based prognostic nomogram is superior to TNM staging. In this heterogeneous patient population, the CLR and CLR-based prognostic model could provide a simplified, cost-effective, widely available, non-invasive, and easily promotive biomarker for tailored prognostic advice.

The expression status of biomarkers (ER, PR, and HER-2) have been shown to be highly correlated with breast cancer outcomes. We have provided a difference and survival analysis for each subtype, as depicted in the attached table and figure. This suggests that CLR may have varying impacts on survival outcomes depending on the subtype of breast cancer. Our findings support the notion that CLR is a promising biomarker for breast cancer prognosis, especially in certain subtypes such as HER2-enriched and luminal B1. However, the lack of statistical significance in other subtypes, such as luminal A and luminal B2, underscores the need for further exploration of the relationship between CLR and survival outcomes in these specific contexts. We propose that in the future, research should delve deeper into the subtype-specific implications of CLR as a biomarker to better understand its potential applications in personalized medicine.

In this study, we examined the novel biomarker CLR in clinical diagnosis and survival prediction. Kaplan-Meier curves and multivariate Cox regression analysis suggested that patients with low CLR had better survival outcomes than those with high CLR. Combining CLR with two other clinicopathological variables (N stage and histopathological type), we established a prognostic nomogram that exhibited excellent predictive performance in the validation and training cohorts. Establishing evidence that CLR is independently associated with survival outcomes in patients with BC is the study's primary strength. We expect that in the future, more studies will investigate the critical significance of the CLR score in BC. [11, 12]

Limitations

Our cohort study, however, has several limitations. First, because it is a retrospective study, it is prone to selection bias. Second, the patient cohorts employed in this study might not accurately represent all of the cancer patients diagnosed and treated at the institution because selection bias could be introduced by including only patients with sufficient measures. Second, the model has not undergone external validation, and our dataset is small. We acknowledge the importance of validating our results in independent cohorts to ensure the reliability and applicability of our conclusions. We plan to extend our research by collaborating with multiple institutions and are currently in discussions with several centers to access their patient data for an external validation phase. This will involve analyzing data from different populations and settings to confirm whether our findings are consistent across diverse groups. Furthermore, a variety of clinical circumstances may have an impact on CLR status. Although our study design prioritizes the assessment of CLR as an independent biomarker that precedes treatment, the treatment type and dosage could modulate CLR levels over time and influence prognosis prediction. Consequently, further multicenter external validation is required. To increase the robustness of our findings, we
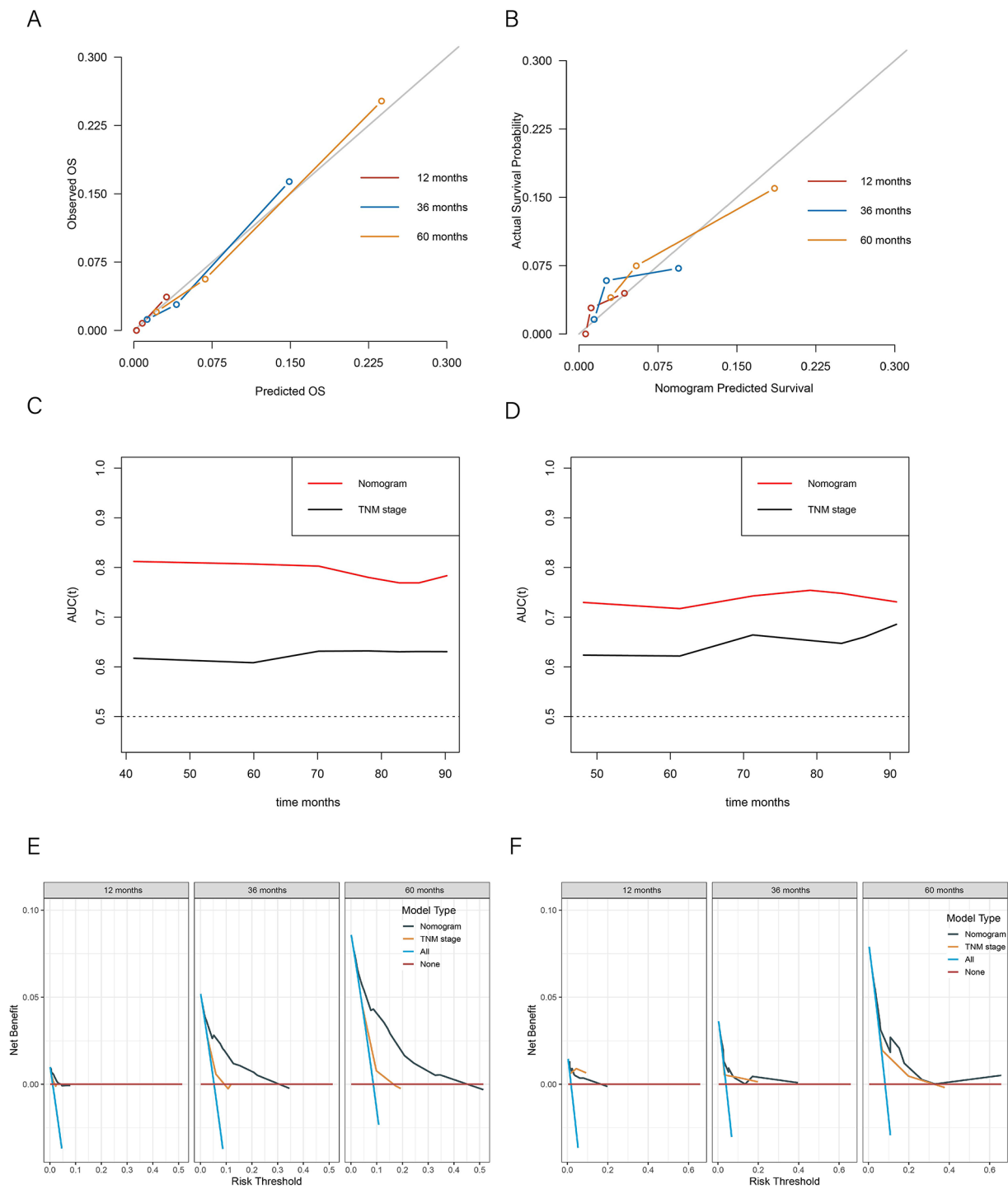
Fig. 4 (See legend on next page.)

intend to gather more data for dynamic analysis. We also intend to incorporate comprehensive treatment data, and suggest prospective research designs to further elucidate the dynamic relationship between CLR, treatment modalities, and survival outcomes.

Conclusion

In patients with BC, the CLR emerges as an independent predictor of OS. Based on CLR and the CLR score, the suggested innovative nomogram model provides a potentially user-friendly tool for individualized prognosis evaluation.

Wan *et al. BMC Cancer*   (2024) 24:914

Page 9 of 10

(See figure on previous page.)

Fig. 4 Evaluation of the prediction efficacy of the prognostic model. (**A**) The survival time predicted by the nomogram was compared with the actual survival time. Nomogram model calibration plot for the training cohort at 1-, 3-, and 5-year. calibration plot, the x-axis typically represents the predicted probabilities, while the y-axis shows the actual fraction of positive cases. Ideally, the plot should show a straight line at a 45-degree angle, indicating perfect calibration where the observed outcomes match the predicted probabilities perfectly. (**B**) The survival time predicted by the nomogram was compared with the actual survival time. Nomogram model calibration plot for the validation cohort at 1-, 3-, and 5-year. (**C**) The new model's predicted accuracy was compared to the training cohort's classical TNM stage using time-independent receiver operating characteristic (ROC) curves. Time-dependent ROC analysis for 1-, 3-, and 5-year OS, the x-axis represents the sensitivity value (true positive rate), and the y-axis represents the 1-specificity (false positive rate). The larger AUC implies that patients can obtain the maximum prediction using this model. (**D**) In the validation cohort, time-independent ROC curves compared the new model's predicted accuracy to the classic TNM stage. (**E**) The training cohort's net benefit rate for the present model and the conventional TNM stage were evaluated using decision curve analysis. The DCA curve is a graphical representation that plots the range of threshold probabilities on the x-axis against the net benefit on the y-axis. Net benefit is calculated as the difference between the proportion of true positives identified (benefits) minus the proportion of false positives identified (harms), weighted by the relative harm of a false positive compared to a false negative. (**F**) The validation cohort's net benefit rate for the present model and the conventional TNM stage were compared using decision curve analysis. *Abbreviations* AUC = area under the curve; OS = overall survival; TNM = tumor-node-metastasis. ROC = receiver operating characteristic

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-024-12648-1.

Supplementary Material 1 Supplement table: Summary descriptives table by groups of 'subtype'.

Supplementary Material 2

Data availability
The data presented in this study are available on request from the corresponding author.

Declarations

Ethical approval
In studies involving human participants, all procedures were conducted in strict adherence to the ethical guidelines set out by the relevant national or institutional research committee, the 1964 Helsinki Declaration, and any subsequent revisions thereof, or similar standards of ethics. The SYSUCC Clinical Research Ethics Committee approved this investigation. All participants were required to provide informed permission before their inclusion in the study.

Competing interests
The authors declare no competing interests.

Author details
¹State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou 510060, China
²Department of Breast Oncology, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China
³Department of Endocrinology, Jiangxi Provincial People's Hospital, The First Affiliated Hospital of Nanchang Medical College, Nanchang 330000, China
⁴Department of Oncology, Guangdong Provincial Hospital of Integrated Traditional Chinese and Western Medicine, Affiliated Nanhai Hospital of Traditional Chinese Medicine of Jinan University, Foshan 528200, China
⁵Department of Radiation Oncology, Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital, Shanghai 200025, China
⁶Department of Medical Oncology, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, China

References
1. Siegel RL, Miller KD, Fuchs HE, Jemal A, Cancer. statistics, 2022. Ca-a Cancer Journal for Clinicians. 2022;72(1):7–33.
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer Stat China 2015 Ca-a Cancer J Clin. 2016;66(2):115–32.
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Ca-a Cancer J Clin. 2021;71(3):209–49.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. Ca-a Cancer J Clin. 2020;70(1):7–30.
5. Yeo SK, Guan J-L. Breast Cancer: multiple subtypes within a Tumor? Trends Cancer. 2017;3(11):753–60.
6. Wang Y, Li Y, Jing Y, Yang Y, Wang H, Ismtula D, et al. Tubulin alpha-1b chain was identified as a prognosis and immune biomarker in pan-cancer combing with experimental validation in breast cancer. Sci Rep. 2024;14(1):8201.
7. Wang Y, Bi X, Luo Z, Wang H, Ismtula D, Guo C, Gelsolin. A comprehensive pan-cancer analysis of potential prognosis, diagnostic, and immune biomarkers. Front Genet. 2023;14:1093163.
8. Schaduangrat N, Homdee N, Shoombuatong W. StackER: a novel SMILES-based stacked approach for the accelerated and efficient discovery of ERα and ERβ antagonists. Sci Rep. 2023;13(1):22994.
9. Ji X, Tian X, Feng S, Zhang L, Wang J, Guo R et al. Intermittent F-actin perturbations by magnetic fields inhibit breast Cancer metastasis. Research (Washington, DC). 2023;6:0080.
10. Zitvogel L, Pietrocola F, Kroemer G. Nutrition, inflammation and cancer. Nat Immunol. 2017;18(8):843–50.
11. Zhou S, He Q, Sheng N, Gong J, Ren J, Wang Z. Blood cholesterol-to-lymphocyte ratio as a novel prognostic marker to predict postoperative overall survival in patients with colorectal cancer. World J Surg Oncol. 2022;20(1):18.
12. Deng Y, Chen Q, Chen J, Zhang Y, Zhao B, Bi X, et al. An elevated preoperative cholesterol-to-lymphocyte ratio predicts unfavourable outcomes in colorectal cancer liver metastasis patients receiving simultaneous resections: a retrospective study. BMC Surg. 2023;23(1):131.

13. Matsubara D, Shoda K, Kubota T, Kosuga T, Konishi H, Shiozaki A, et al. Preoperative total cholesterol-lymphocyte score as a novel immunonutritional predictor of survival in gastric cancer. Langenbecks Arch Surg. 2019;404(6):743–52.
14. Ma Y, Zhang J, Chen X. Lymphocyte-to-monocyte ratio is Associated with the poor prognosis of breast Cancer patients receiving Neoadjuvant Chemotherapy. Cancer Manage Res. 2021;13:1571–80.
15. Huang Z-Z, Song C-G, Huang J-J, Xia W, Bi X-W, Hua X, et al. Prognostic significance of the Controlling Nutritional Status (CONUT) score in surgically treated breast cancer patients. Gland Surg. 2020;9(5):1370–9.
16. Huang Z-Z, Hua X, Song C-G, Xia W, Bi X-W, Yuan Z-Y et al. The Prognostic Prediction Value of systemic inflammation score and the development of a Nomogram for patients with surgically treated breast Cancer. Front Oncol. 2020;10.
17. Hua X, Long Z-Q, Zhang Y-L, Wen W, Guo L, Xia W et al. Prognostic value of preoperative systemic Immune-inflammation index in breast Cancer: a propensity score-matching study. Front Oncol. 2020;10.
18. Caccialanza R, Goldwasser F, Marschal O, Ottery F, Schiefke I, Tilleul P et al. Unmet needs in clinical nutrition in oncology: a multinational analysis of real-world evidence. Therapeutic Adv Med Oncol. 2020;12.
19. Senesse P, Assenat E, Schneider S, Chargari C, Magne N, Azria D, et al. Nutritional support during oncologic treatment of patients with gastrointestinal cancer: who could benefit? Cancer Treat Rev. 2008;34(6):568–75.
20. van der Wal HH, Beverborg NG, Dickstein K, Anker SD, Lang CC, Ng LL, et al. Iron deficiency in worsening heart failure is associated with reduced estimated protein intake, fluid retention, inflammation, and antiplatelet use. Eur Heart J. 2019;40(44):3616–25.
21. Barao K, Vicente Cavagnari MA, Fucuta PS, Forones NM. Association between nutrition status and survival in elderly patients with colorectal cancer. Nutr Clin Pract. 2017;32(5):658–63.
22. Norman K, Pichard C, Lochs H, Pirlich M. Prognostic impact of disease-related malnutrition. Clin Nutr. 2008;27(1):5–15.
23. Kalli S, Semine A, Cohen S, Naber SP, Makim SS, Bahl M. American Joint Committee on Cancer's staging system for breast Cancer, Eighth Edition: what the Radiologist needs to know. radiographics. 2018;38(7):1921–33.
24. Galon J, Mlecnik B, Bindea G, Angell HK, Berger A, Lagorce C, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. J Pathol. 2014;232(2):199–209.
25. Li G, Chen J-Z, Chen S, Lin S-Z, Pan W, Meng Z-W, et al. Development and validation of novel nomograms for predicting the survival of patients after surgical resection of pancreatic ductal adenocarcinoma. Cancer Med. 2020;9(10):3353–70.
26. Miao D-l, Song W, Qian J, Zhu Z-g, Wu Q, Lv C-g, et al. Development and validation of a nomogram for predicting overall survival in pancreatic Neuro-endocrineTumors. Translational Oncol. 2018;11(5):1097–103.