72723_Auto_Edited.docx

Retrospective Study

Added value of systemic inflammation markers for monitoring response to neoadjuvant chemotherapy in breast cancer patients

Ke ZR *et al*. Monitoring response to neoadjuvant chemotherapy

Zi-Rui Ke, Wei Chen, Man-Xiu Li, Shun Wu, Li-Ting Jin, Tie-Jun Wang

Abstract

BACKGROUND

Complete response after neoadjuvant chemotherapy (rNACT) elevates the surgical outcomes of patients with breast cancer, however, non-rNACT have a higher risk of death and recurrence.

AIM

To establish novel machine learning (ML)-based predictive models for predicting probability of rNACT in breast cancer patients who intends to receive NACT.

METHODS

A retrospective analysis of 487 breast cancer patients who underwent mastectomy or breast-conserving surgery and axillary lymph node dissection following neoadjuvant chemotherapy at the Hubei Cancer Hospital between January 1, 2013, and October 1, 2021. The study cohort was divided into internal training and testing datasets in a 70:30 ratio for further analysis. A total of twenty-four variables were included to develop predictive models for rNACT by multiple ML-based algorithms. A feature selection

approach was used to identify optimal predictive factors. These models were evaluated by the receiver operating characteristic (ROC) curve for predictive performance.

RESULTS

Analysis identified several significant differences between the rNACT and non-rNACT groups, including total cholesterol, low-density lipoprotein, neutrophil-to-lymphocyte ratio, body mass index, platelet count, albumin-to-globulin ratio, platelet-to-lymphocyte ratio, and lymphocyte-to-monocyte ratio. The areas under the curve of the six models ranged from 0.81 to 0.96. Some ML-based models performed better than models using conventional statistical methods in both ROC curves. The support vector machine (SVM) model with twelve variables introduced was identified as the best predictive model.

CONCLUSION

By incorporating pretreatment serum lipids and serum inflammation markers, it is feasible to develop ML-based models for the preoperative prediction of rNACT and therefore facilitate the choice of treatment, particularly the SVM, which can improve the prediction of rNACT in patients with breast cancer.

Key Words: Breast cancer; Neoadjuvant chemotherapy; Clinical response; Machine learning; Prediction

Ke ZR, Chen W, Li MX, Wu S, Jin LT, Wang TJ. Added value of systemic inflammation markers for monitoring response to neoadjuvant chemotherapy in breast cancer patients. *World J Clin Cases* 2022; In press

Core Tip: For predicting after neoadjuvant chemotherapy (rNACT), some machine learning-based models performed better than models using conventional methods, and the support vector machine model performed best. Preoperative serum lipids and

serum inflammation markers have contributed to predicting rNACT in breast cancer patients. These results suggested the need to raise awareness of the importance of minimally-invasive approaches for monitoring breast cancer patients who intended to undergo neoadjuvant chemotherapy. However, the current study needs to be validated with caution and require external validation in the future.

INTRODUCTION

Worldwide, breast cancer is a major cause of human suffering and high mortality among women[1]. Neoadjuvant chemotherapy (NACT) as a treatment for early breast cancer, can make breast conserving surgery more feasible, and may achieve more than the same chemotherapy after surgery to eradicate micrometastasis[2]. More than 65% of the patients treated with NACT have a response, and more than 15% have achieved a complete clinical response. Although some trials use the old chemotherapy regimen, more than 15% of the patients have undergone partial chemotherapy[2-4]. In other words, most patients who cannot achieve a complete pathological response after NACT may face a higher risk of death and recurrence. Therefore, it is necessary to develop a practical, convenient and efficient tool to predict the pathological response of patients with NACT breast cancer.

ML-based integrated analysis is a new computer-based method, which has been widely used in medical data management in the past decade[5]. It appears at the intersection of statistics and computer science. The former attempts to learn relationships from data, while the latter emphasizes efficient computational algorithms[6,7]. Compared with traditional statistical prediction models such as logistic regression (LR), ML depends on a predetermined model. It can potentially find the interaction between variables and iteratively learn the update algorithm from the data[8,9]. Previously, several conventional predictive models have been made for predicting after NACT in breast cancer patients, including LR, GLM[10-14]. However, few reports have incorporated multiple ML-based ensemble analyses for predicting after NACT (rNACT).

In this study, we aimed to develop an rNACT risk prediction model for breast cancer patients that utilizes pretreatment serum lipids and serum inflammation markers to stratify patients by rNACT risk on admission. We then analyzed the predictive performance of these ML-based models in a deviation cohort and then verified performance in an internal and external validation cohort.

MATERIALS AND METHODS

Patients

Between January 1, 2013 and October,1, 2021, we retrospectively collated data from consecutive patients who had been diagnosed with breast cancer at the Hubei Cancer Hospital. All patients had received NACT before surgery. This study was approved by the Institutional Ethics Committee of the Hubei Cancer Hospital (Reference: LLHBCH2021YN-021), in compliance with the Declaration of Helsinki. Written informed consent was obtained from all participants before any treatment. We confirmed that the data from all the patients were anonymized in this study. The inclusion and exclusion criteria were summarized in Figure 1. The study cohort was divided into internal training and testing datasets in a 70:30 ratio for further analysis.

Blood data collection

The blood samples of all patients were taken from the fasting state before chemotherapy, and the blood tests were operated by professional personnel to ensure that the blood test results were not biased. The results of the blood test are as follows: Blood routine, liver and kidney function, electrolytes, and blood lipids.

Evaluating the safety and efficacy of NACT

According to RECIST (version 1.1) criteria[15], the efficacy of NACT is defined as follows: (1) Cardiol Res. The tumor is disappeared completely; (2) Partial response (PR). The diameter of the tumor is reduced($\geq$ 30%); (3) Progressive disease (PD). The diameter of the tumor was reduced($\geq$ 20%); (4) Stable disease (SD). The diameter of the tumor was

altered between PR and PD. Collectively, patients were considered to be responsive to NACT provided that they were evaluated as CR or PR after NACT treatment. On the contrary, patients with SD or PD were regarded as non-responsive to NACT.

Development and validation of ML-based models

Four ML-based algorithms were performed to build predictive models, we used the *caret* package to randomly divide the data set into two parts, 70% for model training and 30% for model testing. A total of 6 ML-based algorithms were executed to establish the predictive model. According to the principle of "two-step estimation"[16], we have gradually screened the model variables, as follows:

The characteristic variable is marked as X and the target variable is marked as Y. The X and Y were evenly divided into two parts, namely X1, Y1, and X2, Y2. Through univariate screening, the variable quantum set $\boxed{\times}$ was screened on X1 and Y1, and $\boxed{\times}$ was filtered by X2 and Y2. Then, a lasso was used to fit the model again, and the filtered variables were marked as $\boxed{\times}$ and $\boxed{\times}$. Briefly, by sorting the intersection of variable sets, the optimal subset modeling is obtained. The model was evaluated by inspection, discrimination, and calibration. The receiver operating characteristic (ROC) curve was used to evaluate the recognition ability of the prediction model in the training data set and the test data set; The discrimination ability of each model was quantified by the area under the ROC curve (AUC), decision curve analysis(DCA), and clinical impact curve(CIC).

Statistical analysis

Continuous variables are expressed as mean (standard deviation) and compared using the two-tailed t-test or the Mann-Whitney test. Categorical variables were compared using the chi-square test or Fisher's exact test. Univariate and multivariate logistic analyses were used to explore the risk factors for rNACT. Several ML-based algorithms

were applied to predict rNACT, including support vector machine(SVM), random forest(RF), Naive Bayes(NB), neural network(NN), decision tree(DT), and generalized linear model (GLM)[17,18]. Among all 6 algorithms, the GLM is considered conventional methods, and the others are representative supervised ML-based algorithms. The prediction ability of the 6 models was first evaluated by the ROC curve. All analysis was performed using the Python programming language (version 3.9.2, Python Software Foundation, https://www.python.org/) and R Project for Statistical Computing (version 4.0.4, http://www.r-project.org/). Statistical analyses were performed using a two-tailed Student's t-test in PRISM software (GraphPad 6 Software) to compare the differences between rNACT and non-rNACT groups assuming equal variance. All P values were two-tailed, and $P < 0.05$ was considered statistically significant.

RESULTS

Clinicopathological characteristics

During the period of enrollment, 287 consecutive patients with breast cancer underwent mastectomy or breast-conserving surgery and axillary lymph node dissection following NACT. Besides, 201 patients were validated as external data sets for the prediction model. Demographics and baseline data were summarized in Table 1. According to the RECIST (version 1.1) criteria, rNACT was identified in 255 (88.9%) and 32 (11.1%) patients with non-rNACT in the internal whole cohort. In the external cohort, 176(88.0%) patients were confirmed to have rNACT, and 24 (12.0%) patients represented non-rNACT. Overall, most patients with breast cancer presenting with rNACT were positively associated with pretreatment serum lipids and serum inflammation markers. No statistically significant difference was detected between two cohorts with regard to age, menopause, grade, smoking, estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). ($P \geq 0.05$).

Variable importance and candidate features selection

By feature selection, the twenty-four variables for each algorithm were screened by their predictive importance. As depicted in Figure 2A, only twelve of the candidate features were eventually chosen for modeling, among which eight features had a positive association with rNACT, including PLT, monocyte count (MONO), neutrophil-to-lymphocyte ratio (NLR), lymphocyte-to-monocyte ratio (LMR), platelet-to-lymphocyte ratio (PLR), low-density lipoprotein(LDL), A/G, and total cholesterol (TC). Four features were negatively correlated with rNACT, including high-density lipoprotein (HDL), triglyceride (TG), BMI, and age. The weight of the top eight variables was shown in Figure 2B. The pretreatment serum lipids and serum inflammation markers also showed significant differences between rNACT and non-rNACT groups (Figure 2C-J). Multivariable logistic analysis using raw data of the candidate features proved that the features selected by stepwise analysis exhibited similar risk implications (Supplementary Table 1). Based on our results, NLR(OR: 1.02, 95%CI: 0.78-1.26), LMR(OR: 1.44, 95%CI: 1.32-1.56), PLR(OR:2.54, 95%CI: 1.81-6.94), PLT(OR:1.87, 95%CI: 1.76-1.98), LDL(OR:1.01, 95%CI: 0.89-1.13), BMI(OR:1.23, 95%CI: 0.78-1.68), A/G(OR:1.69, 95%CI: 1.24-2.14), TC(OR:0.71, 95%CI: 0.26-1.16), and TG (OR:0.42, 95%CI: 0.17-0.68) were positively correlated with rNACT.

Comparison Between ML-Based Models

A total of twelve preoperative variables were used to develop predictive models for rNACT based on six algorithms. The predictive performance of all models was shown in Figure 3A and B and Table 2. The best performance was observed in the SVM model (AUC = 0.96, 95%CI: 0.91-1.01), which performed similarly to RF model (AUC = 0.94, 95%CI: 0.87-1.01), superior than NB model (AUC = 0.86, 95%CI: 0.79-0.93), NN model (AUC = 0.88, 95%CI: 0.82-0.94), DT model (AUC = 0.83, 95%CI: 0.77-0.89), and GLM (AUC = 0.81, 95%CI: 0.71-0.91). All ML-based models were better than conventional model. Furthermore, the optimal model SVM showed superior to the traditional linear model in discrimination (Figure 3C and D).

Internal and external validation of the optimal predictive model

To further validate the performance of the SVM model, we also adopted CIC to evaluate the prediction efficiency, as illustrated in Figure 4A, the CIC demonstrated that the stratification of rNACT could be distinguished in the training cohorts. These results were also parallel to risk factors of rNACT delineated in the validation cohorts (Figure 4B), indicating that the selected features were highly relevant to rNACT.

DISCUSSION

Reliable markers of chemosensitivity help select patients who most benefit from NACT[19]. Previous studies on the candidate predictors of NACT efficacy in breast cancer patients are discordant, suggesting that the potential predictors to predict efficacy is insufficient[20-23]. In addition, whilst many studies report the predictive outcomes of breast cancer patients who have received NACT, however, relatively few have investigated the individual contribution of multiple models to accuracy, especially prediction efficiency[24-27]. Whilst this study indicates that ML-based predictive algorithms should be included in NACT risk assessments in breast cancer patients, it also highlights the importance of conducting newly predictive models for clinical management.

Supervised ML algorithms have been a dominant method in the data mining field[17]. In recent years, ML-based algorithms were widely used for the evaluation of disease prognosis[28-30]. In this study, extensive variables were made to identify those predictive that applied more than one supervised ML algorithm on rNACT prediction. Based on the ML algorithm, we employed a variety of statistical, probabilistic, and optimization methods to learn from experience and detect useful patterns from large, unstructured, and complex datasets. To sum up, we extracted the data from the patient's medical records as much as possible. With the help of different algorithms, such as automated text categorisation[31], network intrusion detection[32], optimizing manufacturing process[33], *etc.*, we finally obtained meaningful candidate variables. Given the growing

applicability and effectiveness of supervised ML algorithms on predictive disease modeling. Interestingly, we found that the SVM algorithm is applied most robust in predicting rNACT, which denotes superior performance than the conventional linear prediction model. Besides, the remaining machine prediction models are better than GLM. Therefore, our research demonstrated that, compared with the traditional model, machine learning modeling prediction rNACT could obtain better prediction performance.

Inflammation is associated with the development and malignant progression of most cancers[34]. Inflammatory blood markers have emerged as potential prognostic factors in various cancers，such as NLR, LMR, and PLR. Activated inflammatory cells are sources of reactive oxygen species and reactive nitrogen intermediates that can promote cancer initiation[35]. In breast cancer, pretreatment NLR values are associated with patient prognosis[36]. Similarly, our study indicated that NLR, LMR, and PLR values can be reliably used to predict breast patient responses to NACT treatment, which can effectively stratify patients based upon their likelihood of achieving rNACT. Besides, we also found that pretreatment abnormal A/G ratio, which might be attributable to rNACT. Indeed, a low pretreatment A/G ratio is associated with poor prognosis in human cancers[37]. The importance of lipids in tumor progression, invasion, and metastasis has been described in the previous studies[38]. High triglycerides and low levels of HDL are observed to promote tumor growth[39]. In the present study, we observed that LDL, TC, and BMI were highly associated with rNACT, consistent with previous studies[37,38]. Collectively, clinicians can more effectively weigh the relative costs and benefits of pretreatment serum lipids and serum inflammation markers to ensure that they act in the optimal choice of breast cancer patients.

There are multiple strengths to this study. First, our observations were limited to retrospective studies from a single-center, these findings need further multi-institutional validation with larger sample size. Second, our nomograms were merely validated *via* an internal training set, external verification using independent patient set is necessary. Third, this is a retrospective study that could not completely avoid missing

data and measurement biases, more candidate useful biomarkers may be needed to develop predictive models in the future.

CONCLUSION

In summary, for predicting rNACT, some ML-based models performed better than models using conventional methods, and the SVM model performed best. Preoperative serum lipids and serum inflammation markers have contributed to predicting rNACT in breast cancer patients. These results suggested the need to raise awareness of the importance of minimally-invasive approaches for monitoring breast cancer patients who intended to undergo NACT. However, the current study needs to be validated with caution and require external validation in the future.

72723_Auto_Edited.docx