Postprint: Predicting Neoadjuvant Therapy Efficacy in Breast Cancer Using Deep Learning-Assisted Biopsy Pathological Images

Authors: Yunzhao Luo, Jiang Hongchuan, Xu Feng, Xu Feng

Date: 2024-09-25T00:00:00+00:00

Abstract

Background: Preoperative neoadjuvant therapy (NAT) is the standard treatment for locally advanced breast cancer; however, only a subset of patients are sensitive to NAT. Predicting treatment response prior to NAT is crucial. Previous studies utilizing statistical methods combined with clinical data or deep learning methods combined with radiological images for predicting breast cancer NAT efficacy have yielded suboptimal results. Objective: To develop a deep learning (DL-CNB) model based on breast cancer core needle biopsy whole slide images (WSI) using multiple instance learning (MIL) for predicting pathological complete response (pCR) and visualizing relevant tumor regions. Methods: This retrospective study collected clinical data and pre-NAT core needle biopsy hematoxylin-eosin (H-E) stained sections from breast cancer patients who received neoadjuvant therapy at Beijing Chaoyang Hospital between April 2019 and April 2022. A total of 195 patients were selected according to inclusion and exclusion criteria. Patients were categorized into pCR group (MP=5, n=40) and non-pCR group (MP=1~4, n=155) based on Miller-Payne (MP) grading. Clinical data were first analyzed to construct a Logistic regression model for pCR influencing factors. All WSIs were randomly partitioned into training and test sets at a 4:1 ratio, with 25% of the training set reserved as validation set. All tumor cell regions in each WSI were annotated. The training set was prepared through sliding window patching, data screening, data augmentation, and normalization. Five convolutional neural network models were compared to select the optimal model as the feature extractor for DL-CNB. Parameters were configured to train the DL-CNB model. The model was evaluated using an independent test set to assess the predictive value of DL-CNB. Heatmaps generated from attention module weights were used to visualize important WSI regions relevant to prediction. Results: The pCR group exhibited significantly higher proportions of patients with high histological grade, ER negativity, PR negativity, HER2 positivity, and high Ki-67 expression compared to the non-pCR group (P<0.05). Compared with HR+/HER2-, HR-/HER2+ (OR=10.189, 95%CI=3.225~32.187) and HR+/HER2+ (OR=3.349, 95%CI=1.152~9.737) showed statistically significant associations with pCR achievement (P<0.05). The Logistic regression model achieved an AUC of 0.769 and accuracy of 81.000%. The DL-CNB model achieved an AUC of 0.914 and accuracy of 84.211% on the independent test set. WSI tumor regions with non-pCR and pCR labels from the independent test set were randomly selected for visualization demonstration. Conclusion: The DL-CNB model enables prediction of neoadjuvant therapy pCR and visualization of important regions through breast cancer biopsy WSI, demonstrating superior performance compared to clinical data prediction models. This study can provide clinical decision-making reference for breast cancer patients eligible for NAT, assist in achieving individualized precision treatment, and holds significant importance for improving patient quality of life and survival prognosis.

Full Text

Predicting Response to Neoadjuvant Therapy in Breast Cancer Using Deep Learning on Primary Core Needle Biopsy Slides

LUO Yunzhao, JIANG Hongchuan, XU Feng*

Department of Breast Surgery, Beijing Chaoyang Hospital of Capital Medical University, Beijing 100020, China

*Corresponding author: XU Feng, Associate chief physician; E-mail: drxufeng@mail.ccmu.edu.cn

Abstract

Background: Preoperative neoadjuvant therapy (NAT) is a standardized treatment for locally advanced breast cancer. However, only a portion of patients are sensitive to NAT, making it crucial to predict treatment efficacy before NAT initiation. Previous studies using statistical methods combined with clinical data or deep learning methods combined with medical imaging to predict NAT efficacy in breast cancer have yielded suboptimal results.

Objective: A deep learning model based on core-needle biopsy whole slide images (WSI) of breast cancer (DL-CNB) was trained using the multiple instance learning (MIL) method to predict pathological complete response (pCR) and visualize related tumor areas.

Methods: A retrospective study was conducted to collect clinical data and biopsy hematoxylin-eosin (H-E) stained slides from breast cancer patients who received neoadjuvant therapy at Beijing Chaoyang Hospital from April 2019 to April 2022. A total of 195 patients were selected according to inclusion and exclusion criteria. Patients were divided into pCR group (MP=5, n=40) and non-pCR group (MP=1-4, n=155) based on Miller-Payne (MP) grading. Clinical data were analyzed to construct a Logistic regression model of pCR influencing factors. All WSI images were randomly divided into training and test sets at a 4:1 ratio, with 25% of the training set used as a validation set. All tumor cell regions in each WSI were labeled, and the training set was prepared through sliding window extraction, data screening, data augmentation, and normalization. Five convolutional neural network models were compared, and the optimal model was selected as the feature extractor for DL-CNB. Parameters were set to train the DL-CNB model. The predictive value of DL-CNB was evaluated using an independent test set. To visualize important regions related to prediction in WSI, heat maps were drawn based on weights obtained from the attention-based module.

Results: The proportion of patients with high histological grade, ER negativity, PR negativity, HER2 positivity, and high Ki-67 expression was significantly higher in the pCR group than in the non-pCR group ($P<0.05$). Compared with HR+/HER2-, HR-/HER2+ (OR=10.189, 95%CI=3.225-32.187) and HR+/HER2+ (OR=3.349, 95%CI=1.152-9.737) showed statistically significant associations with achieving pCR ($P<0.05$). The Logistic regression model predicted an AUC of 0.769 with an accuracy of 81.000%. The DL-CNB model achieved an AUC of 0.914 and an accuracy of 84.211% on the independent test set. Representative tumor regions labeled as non-pCR and pCR from the independent test set were randomly selected for visual display.

Conclusion: The DL-CNB model enables prediction of pCR in neoadjuvant therapy and visualization of important regions through breast cancer biopsy WSI. The prediction results were superior to the clinical data Logistic regression method. Therefore, this study can provide clinical decision-making reference for breast cancer patients who meet NAT indications, assist in realizing individualized precision treatment, and is of great significance for improving patients' quality of life and survival expectancy.

Keywords: Neoadjuvant therapy for breast cancer; Biopsy pathological WSI; Deep learning model; Multiple instance learning algorithm; Precision therapy

Introduction

Preoperative neoadjuvant therapy (NAT) is a standardized treatment for locally advanced breast cancer [1]. Patients who achieve pathological complete response (pCR) after NAT have longer overall survival and disease-free survival rates [2]. However, although the significance of neoadjuvant therapy has been recognized, only some patients are sensitive to NAT. Those with stable disease or disease progression after NAT not only endure treatment-related adverse effects but also miss the optimal treatment window, leading to poor prognosis [3]. Therefore, predicting treatment efficacy before NAT is crucial.

In recent years, artificial intelligence has made tremendous progress in medical image processing. Studies have successfully trained deep learning (DL) models using pre-NAT multiparametric ultrasound images [4], contrast-enhanced magnetic resonance imaging (MRI) [5-6], and mammography images [7] to predict pCR after neoadjuvant therapy. Compared with imaging, pathology is the gold standard for breast cancer diagnosis, and core needle biopsy is a mandatory examination for every patient diagnosed with breast cancer. Currently, DL models have been used across various tumors to identify cancer cells in whole slide images (WSI) [8], diagnose diseases [9], and even predict molecular subtypes [10], patient survival rates [11-12], and disease recurrence risk [13] from WSI. However, research on using DL models to learn breast cancer cell characteristics and build NAT efficacy prediction models remains limited, with prediction models trained solely on biopsy pathology images showing suboptimal performance [14]. This study aims to construct a DL model based on tumor cell regions in breast cancer biopsy WSI to predict pCR and visualize related tumor areas.

1.1 Study Subjects

A retrospective analysis was conducted on clinical data and pre-NAT core needle biopsy hematoxylin-eosin (H-E) stained slides from 195 breast cancer patients who received neoadjuvant therapy at Beijing Chaoyang Hospital between April 2019 and April 2022. Inclusion criteria were: (1) pathological diagnosis of invasive breast cancer by core needle biopsy; (2) completion of $ $4 cycles of standardized NAT; (3) surgical treatment after NAT (including modified radical mastectomy, simple mastectomy plus sentinel lymph node biopsy, or breast-conserving surgery); and (4) complete clinical data. Exclusion criteria were: (1) prior treatment before NAT; (2) severe organic disease preventing completion of full NAT cycles and surgery; (3) distant metastasis; (4) biopsy slides from other hospitals; and (5) lost or low-quality pathology slides. This study was approved by the Beijing Chaoyang Hospital Ethics Committee (approval number: 2021-科-102).

1.2 Grouping

Two or more experienced pathologists determined Miller-Payne (MP) grading [15] by comparing pre-NAT biopsy slides with postoperative specimens. MP grade 5 was classified as the pCR group (n=40), and MP grades 1-4 as the non-pCR group (n=155).

1.3 Data Collection

1.3.1 Clinical Data: Clinical data were recorded for all enrolled patients, including age, maximum tumor diameter, cT stage, cN stage, histological grade, axillary lymph node status, estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), proliferation index Ki-67, and molecular subtype. HER2 positivity was defined as immunohistochemistry (IHC) 3+ or IHC 2+ with positive in situ hybridization (ISH) amplification.

1.3.2 Tumor Regions: All breast cancer biopsy H-E stained slides were scanned into WSI. To avoid inter-observer variability, two physicians independently labeled all tumor cell regions in each WSI. Each patient had 1-3 corresponding slides.

1.4 Deep Learning Model Development

The DL-CNB model was developed using Python 3.7 with an attention-based multiple instance learning (MIL) approach [16]. In MIL, each training sample is a “bag” composed of multiple “instances” ; bags are labeled but individual instances are not. The goal of MIL is to predict the label of the entire bag by considering all instances within it. The DL-CNB model structure is shown in Figure 1 [Figure 1: see original paper]. All WSI images were randomly divided into training and test sets at a 4:1 ratio (the test set was independent with no overlap with training data), with 25% of the training set used as a validation set.

1.4.1 Training Data Preparation: (1) Sliding window patching: Numerous non-overlapping square patches (256$×$256 pixels each) were cropped from each tumor region obtained from original WSI using a sliding window approach. (2) Data screening: The proportion of tumor area in each patch was calculated, and patches with larger tumor area proportions were retained for model training. (3) Data augmentation: M bags were randomly constructed within each WSI, with each bag randomly containing N non-repeating patches. Horizontal flipping and rotation were applied to reduce overfitting and increase model robustness. (4) Data normalization: Training data after augmentation were uniformly normalized to reduce errors from staining or scanning variations during acquisition. (5) Label assignment: All bags from pCR group patients were labeled positive, and those from non-pCR group patients were labeled negative.

1.4.2 Feature Extraction: The N patches in each bag were input as a whole into a convolutional neural network (CNN) for feature extraction to obtain an N-dimensional feature vector. AlexNet, VGG19, ResNet101, DenseNet121, and Inception-v3 models were compared on the training set (including validation set) to select the optimal feature extractor for DL-CNB.

1.4.3 Feature Learning: The extracted N-dimensional feature vectors were first processed through a fully connected layer to a fixed dimension, then input into an attention module (comprising 2 fully connected layers and 1 hyperbolic tangent activation function) to obtain weights for each patch. The N weights and corresponding original feature vectors were weighted and summed before being fed into a classifier to output the bag’ s classification probability. After processing by the attention module, importance weight values for each patch were obtained. Heat maps were created based on these weight values to visualize important regions in each WSI related to prediction.

1.4.4 Model Training and Testing: During training, cross entropy [17] was used as the loss function, and stochastic gradient descent (SGD) [18] was used as the optimizer to update model parameters, with a learning rate of 1e-4 and weight decay of 1e-3. A cosine annealing strategy [19] was used to adjust the learning rate. During testing, model outputs from all bags of patients in the independent test set were aggregated to predict NAT efficacy. Receiver operating characteristic (ROC) curves were plotted, and area under the curve (AUC), accuracy (ACC), sensitivity (SENS), and other metrics were calculated to evaluate model performance.

1.5 Statistical Methods

Clinical data were analyzed using SPSS 20.0 software. Categorical data were expressed as constituent ratios or percentages (%) and compared between groups using Pearson ^{2} test or Yates continuity correction. Multivariate analysis was performed using binary Logistic regression. $P<0.05$ was considered statistically significant.

Results

2.1 Clinical Data Analysis

2.1.1 Comparison of Related Factors Between Groups: No statistically significant differences were found between pCR and non-pCR groups in age, cT stage, or cN stage ($P>0.05$). However, the pCR group had significantly higher proportions of patients with high histological grade, ER negativity, PR negativity, HER2 positivity, high Ki-67 expression, HER2-positive molecular subtype, and triple-negative breast cancer (TNBC) compared to the non-pCR group ($P<0.05$) (Table 1).

2.1.2 Multivariate Analysis of Factors Influencing NAT Efficacy: Using NAT efficacy as the dependent variable (non-pCR=0, pCR=1) and histological grade, Ki-67, and molecular subtype as independent variables in binary Logistic regression analysis, only molecular subtype was identified as an influencing factor for predicting NAT efficacy. Compared with HR+/HER2-, HR-/HER2+ (OR=10.189, 95%CI=3.225-32.187) and HR+/HER2+ (OR=3.349, 95%CI=1.152-9.737) showed statistically significant associations with achieving pCR ($P<0.05$) (Table 2). The Logistic regression model incorporating all patients' clinical data (n=195) predicted an AUC of 0.769 with an accuracy of 81.000%.

2.2 DL-CNB Model Results

2.2.1 Comparison of Different Feature Extraction Models: Five commonly used CNN models for image feature extraction were compared: AlexNet, VGG19, ResNet101, DenseNet121, and Inception-v3. Their performance in DL-CNB model training is shown in Table 3 . In the validation set, VGG19 achieved the highest AUC (0.881) and accuracy (90.000%), outperforming other models.

Therefore, VGG19 was selected as the feature extractor for the DL-CNB model in the testing phase.

2.2.2 DL-CNB Model Performance on Test Set: The independent test set was imported into the trained DL-CNB model for testing (Table 3). The model achieved an AUC of 0.914 and an accuracy of 84.211% on the independent test set, surpassing the Logistic regression model built on all patients' clinical data. Additionally, the ROC curve of the DL-CNB model was consistently higher than that of the clinical data prediction model on the independent test set (Figure 2 [Figure 2: see original paper]).

2.2.3 Visualization of Important Regions: Representative tumor regions labeled as non-pCR and pCR from the independent test set were randomly selected for visualization (Figure 3 [Figure 3: see original paper]). In the heat maps, higher weight values (red) indicate greater contribution to prediction results, while lower weight values (blue) indicate lesser contribution.

Discussion

Breast cancer is the most common malignant tumor among women worldwide. Neoadjuvant therapy for breast cancer can shrink tumor lesions, reduce tumor stage, convert inoperable cases to operable ones, increase breast-conserving rates, and improve postoperative survival and quality of life. With the standardization of breast cancer neoadjuvant therapy, the proportion of patients achieving pCR after NAT has increased annually. However, the toxic effects of chemotherapeutic and targeted agents remain, significantly impacting the quality of life of non-pCR patients. Therefore, clinicians need to differentiate patients before NAT, recommending neoadjuvant therapy only for those who are sensitive to NAT to avoid unnecessary toxicity.

Clinical data are the most accessible and analyzable. Numerous studies have used clinical data from breast cancer patients receiving neoadjuvant therapy to predict NAT efficacy, with consistent and mutually corroborating results. Patients with smaller tumor size [20], higher histological grade [21], higher Ki-67 expression [22], greater tumor-infiltrating lymphocyte density [23], triple-negative subtype, and HER2 positivity [24] are more likely to achieve pCR. This study similarly analyzed clinical data from 40 pCR patients and 155 non-pCR patients, using Logistic regression model performance as a baseline for evaluating the DL model. Our results showed no statistical correlation between tumor size and pCR achievement, which is inconsistent with previous research [20]. This discrepancy may be attributed to the use of standardized neoadjuvant chemotherapy and targeted drugs in recent years, enabling patients with higher cT stages to achieve pCR. Histological grade and Ki-67 expression showed statistically significant differences between pCR and non-pCR groups, consistent with previous conclusions [21-22], but had no statistically significant impact on pCR in multivariate Logistic regression. This study found that molecular subtype is an independent factor affecting NAT efficacy, with HER2-positive and triple-negative patients more likely to achieve pCR, consistent with previous research [24]. In the field of artificial intelligence prediction, one study used five machine learning models to analyze the impact of race, age, tumor size, pathological features, molecular subtype, and clinical grade on NAT efficacy in breast cancer patients receiving neoadjuvant therapy, with the random forest (RF) classifier performing best, achieving an AUC of 0.88 [25], higher than the Logistic regression model used in this study (AUC=0.769). However, compared with the vast features hidden in imaging and pathology images, clinical data contain limited information, and their predictive value warrants consideration.

In recent years, deep learning has made significant progress in breast cancer imaging processing, enabling NAT efficacy prediction. SKARPING et al. extracted features from tumor sites and corresponding contralateral breast sites in mammography images (ResNet18), concatenated features, and performed classification, achieving an AUC of 0.71 [7]. Another study used AlexNet to extract features from axial and sagittal contrast-enhanced MRI images, with classification performed by support vector machine (SVM). Sagittal images outperformed axial images in the test set (ACC=0.73, AUC=0.75), and with clinical data assistance, sagittal image prediction reached ACC=0.84 and AUC=0.80 [5]. TALEGHAMAR et al. [4] used multiparametric quantitative ultrasound images for training, with a residual attention network (RAN) as the feature extraction module, achieving ACC=0.88 and AUC=0.86. However, breast multiparametric quantitative ultrasound and contrast-enhanced MRI are expensive and equipment-intensive, limiting their availability in primary hospitals and preventing them from being included in routine breast cancer examinations. Compared with imaging, pathology is the gold standard for breast cancer diagnosis, and core needle biopsy is a mandatory examination for every patient diagnosed with breast cancer. Therefore, this study used pre-NAT core needle biopsy WSI as the research object, and the DL-CNB model test set achieved AUC and ACC exceeding the above imaging-based DL models without requiring clinical data assistance.

Currently, research using DL methods to learn breast cancer tumor cell characteristics and build NAT efficacy prediction models is limited. Domestic scholars have used the UNet++ network to segment tumor regions and cell nuclei in pre-NAT biopsy WSI from breast cancer patients. Through manual feature extraction and cell nucleus-level feature ranking, the model trained with a classifier could predict MP grading with an accuracy of 82.35% [26]. However, manual feature extraction requires converting image features into computer-analyzable data through human intervention, making it difficult to fully utilize hidden invisible features in images. Subsequent scholars used the Inception V3 network to achieve tumor region segmentation and pCR prediction, employing a fast ensemble deep learning strategy. The prediction model trained solely on biopsy pathology images achieved an AUC of 0.82, inferior to the model combining images and clinical data (AUC=0.89) [14]. This indicates that DL models for learning WSI cell nucleus features and predicting pCR need optimization. This study proposes an attention-based multiple instance learning (MIL) method comprising a feature extraction module and a multiple instance learning module. The feature extraction module found that VGG19 outperformed AlexNet, ResNet, Inception-v3, and other feature extraction networks used in previous breast cancer NAT efficacy prediction models. The constructed DL-CNB model demonstrated excellent performance on the independent test set, with prediction accuracy reaching 84.211% and AUC reaching 0.914, surpassing the aforementioned previous imaging prediction models. Additionally, the attention mechanism is a resource allocation scheme for solving information overload problems, assigning intuitive weight values to instances (patches) in bags, determining which instances are more important for bag labels. This study thus achieved interpretability of prediction results by visualizing important regions in WSI.

This study has several limitations: (1) Data were from a single hospital, lacking multi-center validation. (2) Pathological evaluation of NAT efficacy in our hospital was limited to primary breast lesions without considering axillary lymph node status. Future studies could replace MP grading with residual cancer burden (RCB) grading to strictly define pCR. (3) Tumor region delineation during dataset preparation was time-consuming and labor-intensive; future research could consider using UNet++ networks for automatic tumor region segmentation.

In summary, this study successfully trained a deep learning (DL-CNB) model using an attention-based multiple instance learning method, achieving prediction of pCR in neoadjuvant therapy through breast cancer biopsy WSI. The prediction results outperformed clinical data prediction models. Additionally, this study achieved interpretability of prediction results by visualizing important regions in WSI through heat maps. Through the DL-CNB model, this study can provide clinical decision-making reference for breast cancer patients who meet NAT indications, assist in realizing individualized precision treatment, and is of great significance for improving patients' quality of life and survival expectancy.

Author Contributions: LUO Yunzhao was responsible for collecting and labeling clinical data and biopsy pathology WSI, statistical analysis, building and testing the deep learning model, and writing the initial manuscript. JIANG Hongchuan proposed clinical data research indicators, established inclusion and exclusion criteria, and was responsible for selecting study subjects. XU Feng proposed the research idea, designed the study protocol, was responsible for quality control and review, and takes responsibility for the manuscript. All authors confirmed the final manuscript.

Conflict of Interest: The authors declare no conflict of interest.

ORCID IDs: - LUO Yunzhao: https://orcid.org/0000-0002-8311-2432 - JIANG Hongchuan: https://orcid.org/0000-0001-8602-6049 - XU Feng: https://orcid.org/0000-0002-5709-0188

References

[1] GRADISHAR W J, MORAN M S, ABRAHAM J, et al. Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology[J]. J Natl Compr Canc Netw, 2022, 20(6): 691-722. DOI:10.6004/jnccn.2022.0030.

[2] SPRING L M, FELL G, ARFE A, et al. Pathologic complete response after neoadjuvant chemotherapy and impact on breast cancer recurrence and survival: A comprehensive Meta-analysis[J]. Clin Cancer Res, 2020, 26(12): 2838-2848. DOI:10.1158/1078-0432.Ccr-19-3492.

[3] AGARWAL R, UNNIKRISHNAN U G, KEECHILAT P, et al. Pathological complete response in locally advanced breast cancer after neoadjuvant chemotherapy: survival outcome and its relevance as a surrogate end point[J]. South Asian Journal of Cancer, 2020, 9(3): 136-140. DOI:10.1055/s-0040-1721238.

[4] TALEGHAMAR H, ALI JALALIFAR S, CZARNOTA G J, et al. Deep learning of quantitative ultrasound multi-parametric images at pre-treatment to predict breast cancer response to chemotherapy[J]. Sci Rep, 2022, 12(1): 2244. DOI:10.1038/s41598-022-06171-1.

[5] MASSAFRA R, COMES M C, BOVE S, et al. Robustness evaluation of a deep learning model on sagittal and axial breast DCE-MRIs to predict pathological complete response to neoadjuvant chemotherapy[J]. J Pers Med, 2022, 12(6): 953. DOI:10.3390/jpm12060953.

[6] DAMMU H, REN T, DUONG T Q. Deep learning prediction of pathological complete response, residual cancer burden, and progression-free survival in breast cancer patients[J]. PLoS One, 2023, 18(1): e0280148. DOI:10.1371/journal.pone.0280148.

[7] SKARPING I, LARSSON M, FÖRNVIK D. Analysis of mammograms using artificial intelligence to predict response to neoadjuvant chemotherapy in breast cancer patients: Proof of concept[J]. Eur Radiol, 2022, 32(5): 3131-3141. DOI:10.1007/s00330-021-08306-w.

[8] CAMPANELLA G, HANNA M G, GENESLAW L, et al. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images[J]. Nat Med, 2019, 25(8): 1301-1309. DOI:10.1038/s41591-019-0508-1.

[9] MERCAN E, MEHTA S, BARTLETT J, et al. Assessment of machine learning of breast pathology structures for automated differentiation of breast cancer and high-risk proliferative lesions[J]. JAMA Network Open, 2019, 2(8): e198777. DOI:10.1001/jamanetworkopen.2019.8777.

[10] WOERL A C, ECKSTEIN M, GEIGER J, et al. Deep learning predicts molecular subtype of muscle-invasive bladder cancer from conventional histopathological slides[J]. Eur Urol, 2020, 78(2): 256-264. DOI:10.1016/j.eururo.2020.04.023.

[11] SKREDE O J, DE RAEDT S, KLEPPE A, et al. Deep learning for prediction of colorectal cancer outcome: A discovery and validation study[J]. Lancet, 2020, 395(10221): 350-360. DOI:10.1016/s0140-6736(19)32998-8.

[12] ZHAO K, LI Z H, YAO S, et al. Artificial intelligence quantified tumour-stroma ratio is an independent predictor for overall survival in resectable colorectal cancer[J]. EBioMedicine, 2020, 61: 103054. DOI:10.1016/j.ebiom.2020.103054.

[13] LUCAS M, JANSEN I, VAN LEEUWEN T G, et al. Deep learning-based recurrence prediction in patients with non-muscle-invasive bladder cancer[J]. Eur Urol Focus, 2022, 8(1): 165-172. DOI:10.1016/j.euf.2020.12.008.

[14] LI F L, YANG Y Q, WEI Y N, et al. Deep learning-based predictive biomarker of pathological complete response to neoadjuvant chemotherapy from histological images in breast cancer[J]. J Transl Med, 2021, 19(1): 348. DOI:10.1186/s12967-021-03020-z.

[15] OGSTON K N, MILLER I D, PAYNE S, et al. A new histological grading system to assess response of breast cancers to primary chemotherapy: Prognostic significance and survival[J]. Breast (Edinburgh, Scotland), 2003, 12(5): 320-327. DOI:10.1016/s0960-9776(03)00106-1.

[16] ILSE M, TOMCZAK J M, WELLING M. Attention-based deep multiple instance learning[Z]. International conference on machine learning, 2018: 2127-2136.

[17] WU Y N. Cross Entropy[M]. Computer Vision: A Reference Guide. Boston: Springer International Publishing, 2021: 225-226.

[18] THEODORIDIS S. Stochastic Gradient Descent[M]. Machine learning. Oxford: Academic Press, 2015: 161-231.

[19] LOSHCHILOV I, HUTTER F. SGDR: Stochastic gradient descent with warm restarts[Z]. International Conference on Learning Representations, 2017: 1-16.

[20] GOORTS B, VAN NIJNATTEN T J A, DE MUNCK L, et al. Clinical tumor stage is the most important predictor of pathological complete response rate after neoadjuvant chemotherapy in breast cancer patients[J]. Breast Cancer Res Treat, 2017, 163(1): 83-91. DOI:10.1007/s10549-017-4155-2.

[21] XIAO Y, DING J H, MA D C, et al. Predicting pathological complete response in neoadjuvant dual blockade with trastuzumab and pertuzumab in HER2 gene amplified breast cancer[J]. Front Immunol, 2022, 13: 877825. DOI:10.3389/fimmu.2022.877825.

[22] CHEN P X, WANG C, LU R L, et al. Multivariable models based on baseline imaging features and clinicopathological characteristics to predict breast pathologic response after neoadjuvant chemotherapy in patients with

breast cancer[J]. Breast Care (Basel, Switzerland), 2022, 17(3): 306-315. DOI:10.1159/000521638.

[23] DENKERT C, VON MINCKWITZ G, DARB-ESFAHANI S, et al. Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: A pooled analysis of 3771 patients treated with neoadjuvant therapy[J]. Lancet Oncol, 2018, 19(1): 40-50. DOI:10.1016/s1470-2045(17)30904-x.

[24] HAQUE W, VERMA V, HATCH S, et al. Response rates and pathological complete response by breast cancer molecular subtype following neoadjuvant chemotherapy[J]. Breast Cancer Res Treat, 2018, 170(3): 559-567. DOI:10.1007/s10549-018-4790-2.

[25] METI N, SAEDNIA K, LAGREE A, et al. Machine learning frameworks to predict neoadjuvant chemotherapy response in breast cancer using clinical and pathological features[J]. JCO Clin Cancer Inform, 2021, 5: 66-80. DOI:10.1200/cci.20.00078.

[26] XU CY, XIE JW, YANG CX, et al. Prediction of neoadjuvant chemotherapy efficacy in breast cancer based on histomorphological analysis of pathological puncture sections[J]. Journal of Sichuan University (Medical Science Edition), 2021, 52(2): 279-285. DOI:10.12182/20210360505.

(Received: August 25, 2023; Revised: March 18, 2024) (Editor: LI Tingting)

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.