EBioMedicine 60 (2020) 103018

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Deep learning radiomics of ultrasonography: Identifying the risk of axillary non-sentinel lymph node involvement in primary breast cancer

Xu Guo[a,b,1], Zhenyu Liu[b,e,1], Caixia Sun[b,c,f,1], Lei Zhang[a], Ying Wang[g], Ziyao Li[a], Jiaxin Shi[a], Tong Wu[a], Hao Cui[a], Jing Zhang[h], Jie Tian[b,c,d,e,f,**], Jiawei Tian[a,*]

[a] Department of Ultrasound, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China
[b] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[c] Beijing Advanced Innovation Centre for Big Data-Based Precision Medicine, School of Medicine and Engineering, Beihang University, Beijing, China
[d] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shanxi, China
[e] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China
[f] Key Laboratory of Big Data-Based Precision Medicine, Ministry of Industry and Information Technology, Beihang University, Beijing, China
[g] Department of general surgery, The Second Hospital of Hebei Medical University, Shijiazhuang, Hebei, China.
[h] Department of MRI Diagnosis, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China

ARTICLE INFO

Article History:
Received 27 May 2020
Revised 7 September 2020
Accepted 8 September 2020
Available online xxx

Keywords:
Deep learning radiomics
Ultrasonography
Primary breast cancer
Axillary management
NSLN metastasis in the axilla

ABSTRACT

Background: Completion axillary lymph node dissection is overtreatment for patients with sentinel lymph node (SLN) metastasis in whom the metastatic risk of residual non-SLN (NSLN) is low. However, the National Comprehensive Cancer Network panel posits that none of the previous studies has successfully identified such subset patients. Here, we develop a multicentre deep learning radiomics of ultrasonography model (DLRU) to predict the risk of SLN and NSLN metastasis.

Methods: In total, 937 eligible breast cancer patients with ultrasound images were enrolled from two hospitals as the training set (n = 542) and independent test set (n = 395) respectively. Using the images, we developed and validated a prediction model combined with deep learning radiomics and axillary ultrasound to sequentially identify the metastatic risk of SLN and NSLN, thereby, classifying patients to relevant axillary management groups.

Findings: In the test set, the DLRU yields the best performance in identifying patients with metastatic disease in SLNs (sensitivity=98.4%, 95% CI 96.6−100) and NSLNs (sensitivity=98.4%, 95% CI 95.6−99.9). The DLRU also accurately stratifies patients without metastasis in SLN or NSLN into the corresponding low-risk (LR)-SLN and high-risk (HR)-SLN&LR-NSLN category with the negative predictive value of 97% (95% CI 94.2−100) and 91.7% (95% CI 88.8−97.9), respectively. Moreover, compared with the current clinical management, DLRU appropriately assigned 51% (39.6%/77.4%) of overtreated patients in the entire study cohort into the LR group, perhaps avoiding overtreatment.

Interpretation: The performance of the DLRU indicates that it may offer a simple preoperative tool to promote personalized axillary management of breast cancer.

Funding: The National Nature Science Foundation of China; The National Outstanding Youth Science Fund Project of National Natural Science Foundation of China; The Scientific research project of Heilongjiang Health Committee; The Postgraduate Research &Practice Innovation Program of Harbin Medical University.

Introduction

Since the sentinel lymph node (SLN) was first noted in published studies in the 1920s [1], axillary treatment has evolved from the routine completion of axillary lymph node dissection (ALND) for most breast cancer patients to a selective approach based on the assessment of the SLN by sentinel lymph node dissection (SLND). The general procedure is ALND for all cases with SLN metastasis or when the

* Correspondence to: Jiawei Tian, MD, Department of Ultrasound, The Second Affiliated Hospital of Harbin Medical University, No.246 Xuefu Road, Nan Gang Dist., Harbin, Heilongjiang, China.
** Corresponding author at: Jie Tian, PhD, CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, No.95 Zhongguancun East road, Beijing, China.
E-mail addresses: jie.tian@ia.ac.cn (J. Tian), jwtian2004@163.com (J. Tian).
[1] These authors contributed equally to this work as co-first authors.

https://doi.org/10.1016/j.ebiom.2020.103018

Research in context

Evidence before this study

Up to 73% of patients with sentinel lymph node (SLN) metastatic disease in whom the risk of additional non-sentinel lymph node (NSLN) disease is low have received unnecessary further axillary lymph node dissection (ALND). However, the National Comprehensive Cancer Network panel posits that none of the previous clinical methods has successfully identified such subset patients. We searched PubMed and Web of Science, for research articles with the following terms: "(non-sentinel lymph node OR non-SLN OR NSLN) AND (deep learning OR radiomics OR deep learning radiomics) AND (identify OR identification OR predict OR prediction) AND (ultrasound OR ultrasonography OR ultrasonic OR US) AND (metastasis OR involvement OR invasion) AND breast cancer" with no language restrictions. As yet, there was no known study established the deep learning radiomics model for identifying the risk of SLN and NSLN metastatic disease in breast cancer.

Added value of this study

Completion ALND is overtreatment for breast cancer patients with histologically positive SLN but a low risk of NSLN metastasis. By analysing 3049 ultrasound images, this study developed a prediction model based on deep learning radiomics for risk-assessment of the metastatic disease in SLN and NSLN. In different cohorts, the DLRU yields the favourable performance in identifying the risk of SLN and NSLN metastasis. It demonstrates the promising potential of the DLRU, which may help some breast cancer patients with positive SLN but consistently negative NSLN to avoid overtreatment and promote personalized axillary management. Furthermore, the service conditions of the DLRU is simplified by using the ultrasonography, which is the most convenient imaging. By this means, the DLRU could probably be helpful even for hospitals in resource-limited areas.

Implications of all the available evidence

By detecting the detailed metastatic risk of axillary, the DLRU may offer a preoperative tool to stratify breast cancer patients to appropriate axillary surgery, providing a reference for improving axillary management. Therefore, further testing on a larger data set is warranted to make our model applicable.

identification of SLNs fails. It is worth mentioning that there are postoperative complications and morbidities in varying degrees, whether using SLND or ALND [2,3].

In previous studies, up to 73% of the patients undergoing subsequent ALND for a positive pathology result of SLND, resulting in overtreatment, as SLN was found to be the sole site of the axillary lymph node (ALN) metastasis [4–7]. The National Comprehensive Cancer Network's (NCCN) Panel has recommended that completion ALND could be avoided for the patient with low-risk (LR) of residual non-SLN (NSLN) metastatic disease even though SLN is positive [8]. To identify such patients, the American College of Surgeons Oncology Group (ACOSOG) Z0011 trial and the International Breast Cancer Study Group (IBCSG) 23–01 trial concluded that the risk of NSLN involvement is very low in patients with limited SLN metastasis such that SLND alone can achieve the same survival rate as ALND while avoiding morbidities [2,6,9–11]. Nevertheless, a vivid discussion continues with arguments that the results of the ACOSOG Z0011 trial can only be suitable for patients undergoing breast-conserving surgery followed by whole-breast radiation, and the standard of pathological examination for the limited SLN metastasis in the IBCSG 23–01 trial is extremely exacting for hospitals in resource-limited areas [6,12], hindering the generalization of the experimental conclusions. Moreover, with a rise in screening awareness as well as improvements in the diagnostic method of breast cancer, SLN is free of tumour burden in some patients with an early diagnosis [13], indicating that SLND is unnecessary for them. However, the SLN status can only be obtained by the invasive SLND or SLN biopsy.

In clinics, various methods, such as mammography, ultrasonography (US) and MRI, are used to diagnose axillary metastasis. Because of being radiation-free, favourable repeatability and being easily incorporated into the preoperative breast examinations [14–16], axillary ultrasonography (AUS) has been routinely used to assess the ALN metastatic disease. However, AUS relies on the perception of the radiologist, like other methods, it is unable to distinguish between SLN and NSLN. Since the 21st century, several clinical predictive methods based on the pathological characters of the breast tumour and SLN, such as Memorial Sloan-Kettering Cancer Centre (MSKCC) nomogram, Mayo nomogram, the Tenon score system, and recursive partitioning tools, have been devised to predict the risk of further axillary NSLN involvement to omit over-treated ALND [4,17–21]. However, the latest NCCN guidelines posit that none of these studies successfully identified the LR group of patients with positive SLNs but consistently negative NSLNs [8]. These realities highlight the lack of an available method for preoperative identification of the status of SLN and NSLN in the axilla, which is vital to axillary management.

In recent years, a few studies have indicated the potential value of quantitative radiomic features from medical images to predict the ALN status [22,23]. Nevertheless, these radiomic methods rely on precise tumour boundaries labelled by radiologists and remain incapable of predicting the NSLN of the ALN. Superior to the conventional radiomics, deep learning radiomics (DLR) is a prospective method that automatically learns feature representations, quantifies information from images and has been shown to match and even surpass human performance in addressing the challenges across the spectrum of cancer detection, treatment, and monitoring [24–26]. Further studies have caused a rapid rise in the potential for the application of DLR in breast cancer imaging such as risk assessment, prediction of prognosis, response to therapy, and even distinguishing the number of metastatic ALNs [24,27–29]. Therefore, we hypothesize that DLR can show a great clinical utility in tackling the challenge of predicting the detailed ALN status.

In this study, we construct a multicentre model by combining DLR with the US, the most common and simplest examination for breast cancer, named DLRU. To better meet the clinical requirements, the fundamental purpose of the DLRU is to identify as many high-risk (HR) patients as possible to ensure an appropriately reduced in overtreatment without adverse impacts on survival. Ultimately, the performance of the DLRU indicates that it may help the detection of SLN and NSLN metastasis to customize personalized axillary surgical strategies to avoid overtreatment for some of the patients without SLN or NSLN metastasis.

Materials and methods

Patients and experimental design

The ethical board of Harbin Medical University (HMU) approved this retrospective multicentre study. Written informed consents from patients were waived owing to the retrospective study design, non-invasive nature of the intervention. We obtained verbal consent for the use of their data in this study, and the ethical board waived the need to document the verbal consent (approval number: KY2016–127).

From March 2017 to July 2019, 1576 consecutive women from two affiliated hospitals of HMU (China) came to the hospital for suspicious breast lesions and initially agreed to participate in the study. Amongst these patients, 1280 patients met the inclusion criteria and

considered eligible for this study. After excluding 343 patients, 937 primary breast cancer patients with 3049 US images were enrolled for analysis (details in Supplementary material and Fig. S1), including 542 women from the 2nd affiliated hospital of HMU as the training set and 395 patients from the 3rd affiliated hospital of HMU as the test set. Inclusion and exclusion criteria are indicated in the Supplementary material. The clinicopathological characteristics needed for the study were collected retrospectively after surgery (Table 1).

The detailed status of ALN was confirmed by combining the results of SLND and ALND as the reference, which is the accepted gold standard in clinical work: negative SLN (SLN-), positive SLN but negative NSLN (SLN+&NSLN-), positive SLN, and NSLN (SLN+&NSLN+)

Table 1
Clinical characteristics of patients in the training and test sets.

	Training set (n = 542)			Test set (n = 395)			P^*
	SLN+ (180,33.2%)	SLN- (362,66.8%)	P	SLN+ (185,46.8%)	SLN- (210,53.2%)	P	
Age, years			0.683			0.077	0.524
Mean	52.1	52.5		51.0	52.7		
SD	10.9	10.5		9.7	10.1		
Median	51	52		51	52		
Range	24−79	28−86		26−83	23−78		
Menopausal status			0.960			0.073	0.596
Premenopausal, no. (%)	73 (40.6)	146 (40.3)		87 (47.0)	80 (38.1)		
postmenopausal, no. (%)	107 (59.4)	216 (59.7)		98 (53.0)	130 (61.9)		
Clinical T stage			<0.001			<0.001	<0.001
T1 (≤20 mm), no. (%)	78 (43.3)	248 (68.5)		34 (18.4)	89 (42.4)		
T2 (21 mm −50 mm), no. (%)	96 (53.3)	110 (30.4)		123 (66.5)	104 (49.5)		
T3 (>50 mm), no. (%)	6 (3.3)	44 (6.7)		28 (15.1)	17 (8.1)		
Histologic type			0.851			0.051	0.708
Ductal carcinoma, no. (%)	156 (86.7)	312 (86.2)		156 (84.3)	184 (87.6)		
Lobular carcinoma, no. (%)	6 (3.3)	11 (3.0)		16 (8.6)	5 (2.4)		
Mixed, no. (%)	10 (5.6)	17 (4.7)		9 (4.9)	7 (3.3)		
Other, no. (%)	8 (4.4)	22 (6.1)		4 (2.2)	14 (6.7)		
ER			0.167			0.057	<0.001
positive, no. (%)	129 (71.7)	234 (64.6)		76 (41.1)	99 (47.1)		
negative, no. (%)	46 (25.6)	111 (30.7)		80 (43.2)	68 (32.4)		
Missing, no. (%)	5 (2.8)	17 (4.7)		29 (15.7)	43 (20.5)		
PR			0.057			0.057	<0.05
positive, no. (%)	117 (65.0)	198 (54.7)		78 (42.2)	91 (43.3)		
negative, no. (%)	58 (32.2)	142 (39.2)		78 (42.2)	75 (35.7)		
Missing, no. (%)	5 (2.8)	22 (6.1)		29 (15.7)	44 (21.0)		
HER-2			0.057			0.524	<0.05
positive, no. (%)	93 (51.7)	164 (45.3)		65 (35.1)	63 (30.0)		
negative, no. (%)	82 (45.6)	176 (48.6)		91 (49.2)	102 (48.6)		
Missing, no. (%)	5 (2.8)	22 (6.1)		29 (15.7)	45 (21.4)		
Ki-67			0.063			0.445	<0.001
<20%, no. (%)	64 (35.6)	153 (42.3)		84 (45.4)	97 (46.2)		
>20%, no. (%)	111 (61.7)	185 (51.1)		70 (37.8)	68 (32.4)		
missing, no. (%)	5 (2.8)	24 (6.6)		31 (16.8)	45 (21.4)		
biological subtype			<0.001			<0.05	0.034
HER2 negative/HR positive, no. (%)	137 (76.1)	249 (68.8)		106 (57.3)	132 (62.9)		
Triple negative, no. (%)	14 (7.8)	45 (12.4)		16 (8.6)	13 (6.2)		
HER2 positive, no. (%)	25 (13.9)	59 (16.3)		34 (18.4)	20 (9.5)		
Missing, no. (%)	4 (2.2)	9 (2.5)		29 (15.7)	45 (21.4)		
NSLN status (SLN positive)							<0.001
NSLN positive, no. (%)	83 (46.1)	0		128 (69.1)	0		
NSLN negative, no. (%)	97 (53.8)	362 (100)		57 (30.9)	210 (100)		
US BI-RADS grade			<0.001			<0.001	<0.001
3, no. (%)	4 (2.2)	15 (4.2)		2 (1.1)	3 (1.4)		
4a, no. (%)	5 (5.0)	76 (21.0)		2 (1.1)	13 (6.2)		
4b, no. (%)	29 (16.1)	123 (34.0)		19 (10.3)	41 (19.5)		
4c, no. (%)	41 (22.8)	127 (35.1)		61 (33.0)	128 (61.0)		
5, no. (%)	97 (53.9)	21 (5.8)		101 (54.6)	25 (11.9)		
Location			0.644			0.051	0.087
UOQ, no. (%)	112 (62.2)	215 (59.4)		114 (61.6)	102 (48.6)		
LOQ, no. (%)	28 (15.6)	46 (12.7)		25 (13.5)	30 (14.3)		
UIQ, no. (%)	30 (16.7)	78 (21.5)		27 (14.6)	59 (28.1)		
LIQ, no. (%)	7 (3.9)	15 (4.1)		8 (4.3)	10 (4.8)		
Centre, no. (%)	3 (1.7)	8 (2.2)		11 (5.9)	9 (4.3)		
AUS report							0.001
suspicious, no. (%)	121 (67.2)	16 (4.4)		130 (70.3)	11 (5.2)		
Unsuspicious, no. (%)	59 (32.8)	346 (95.6)		55 (29.7)	199 (94.8)		

Note: ***P*** represents the difference between each clinicopathological variable between the positive SLN and the negative SLN sets. ***P*** * represents the difference between each clinicopathological variable between the training and the test sets.
The chi-square test was used to compare the difference in categorical variables (all variables except age), while a Student's t-test was used to compare the difference in age.
Abbreviations: ER, oestrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; HR, hormone receptor; ALN, axillary lymph node; SLN, sentinel lymph node; NSLN, non-sentinel lymph node; the US, ultrasonography; UOQ, upper outer quadrant; LOQ, lower outer quadrant; UIQ, upper inner quadrant; LIQ, lower inner quadrant; AUS, axillary ultrasonography.

(Table. S1, details in the Supplementary material). The ideal axillary management for breast cancer is based on the status above: (i) SLN-, any kind of surgery is redundant; (ii) SLN+&NSLN-, SLND is needed; (iii) SLN+&NSLN+, required complete ALND. In this study, we developed the model using DLR to stratify patients into the ideal management groups, including LR-SLN, HR-SLN&LR-NSLN, and HR-SLN&HR-NSLN, corresponding to SLN-, SLN+&NSLN-, and SLN+&NSLN+, respectively. These may help customize the axillary operation.

Ultrasonography of breast lesion and ALN

Before participating in the study, every breast radiologist had to fulfil the US imaging quality control criteria to perform standard breast ultrasound and AUS examinations for patients [30]. To keep the consistency of the imaging, six radiologists with over five years of experience were selected for the preoperative breast and AUS image acquisition (details in Supplementary material), using HITACHI Vision 500 system (Hitachi Medical System, Tokyo, Japan) and Aixplorer US imaging system (SuperSonic Imagine, SSI, France) equipped with a linear probe of 5−13 MHz.

For the breast lesion, 1 or 2 two-dimensional grey-scale static US images at the radial and anti-radial planes from each patient were selected by two radiologists with over ten years of experience. All patients underwent an AUS to assess and record the morphologic appearance of the ALNs in real-time, and the AUS images were subsequently reviewed by the two experienced radiologists. According to the BI-RADS criteria (Supplementary material) [31], the ALN was categorized into “suspicious” or “unsuspicious” as the result of the AUS report (Fig. 1).

Deep learning radiomics model

DLR is a multi-layer computational model designed to extract useful information from image data. The key computational operations are convolution, activation, and pooling. For reducing the risk of overfitting, we introduce a batch of normalization and data augmentation modules. The detailed mathematical description of these operations is illustrated in Supplementary methods.

Two DLR models were designed to predict the metastasis of the SLN and the NSLN, respectively. The DLR models used the same network architecture and training strategy. The structure of the DLR network is shown in Fig. S2. When applying the DLR model, a square region of interest (ROI) containing the entire the full tumour region was manually selected on the US image. Detailly, it contains the entire hypoechoic tumour area, echogenic halo (if present) and some surrounding tissues (which are hyper-echogenic compared with the tumour, such as glands or fatty). If there are posterior and lateral acoustic shadows of the tumour on the US image, the ROI also needs to include part of it. The tumour was not required to be precisely in the centre of ROI. Then, the ROI was resized to 224 × 224 pixels and input to each DLR model. After sequential activation of the convolution and pooling layers, the DLR model predicted the lymph node metastatic probability of the image. For further analysing the information of different orientations of US images, two kinds of images were fed into the two sub DLR models (radial and anti-radial), and the results of the two were fused to give the final metastatic probability of the DLR model (details in Supplementary methods).

Before training this model, we pre-trained the network using 1.28 million natural images in the ImageNet dataset [32], which has shown good performance in disease diagnosis [33,34]. Finally, the weights of the DLR model were fine-tuned by the US images from the breast tumour in the training set.

Integration of the DLRU for identifying the metastatic risk in SLN and NSLN

Based on the DLR models, we developed two sub-models, DLR-1 and DLR-2, which were combined into the DLRU and can sequentially predict the metastatic risk of SLN and NSLN (Fig. 2).

The DLR-1 was constructed using a decision tree (Fig. S3), combining the DLR model and the AUS report (DLR_{+AUS}) to acquire the risk of SLN involvement. Firstly, the US images in two orientations (880 radial and 908 anti-radial) from all breast cancer patients in the training set were used for training the DLR model respectively. Secondly, since the AUS is incapable of distinguishing the SLN from the NSLN, and the lymphatic spread within the axilla is orderly, we assume that the AUS report reflects the metastatic risk of SLN that receives initial drainage from cancer [7,35,36]. Then we extracted the AUS report as a categorical variable with a specific value of 1 (suspicious) or 0 (unsuspicious) and incorporated it into DLR-1. Therefore, in DLR-1, when the DLR model considered the risk of SLN involvement as LR, the AUS report would be subsequently employed to assess the metastatic risk of SLN. If the AUS report considered the SLN as “suspicious”, the metastatic risk would be output as HR. And LR would be output when SLN was considered as “unsuspicious” by AUS report. When the DLR model considered the metastatic risk of SLN is high, that will be directly output. Finally, patients identified as LR-SLN are directly exported by DLRU, and HR-SLN patients will be subsequently input to DLR-2 to continue to identify the risk of NSLN metastasis.

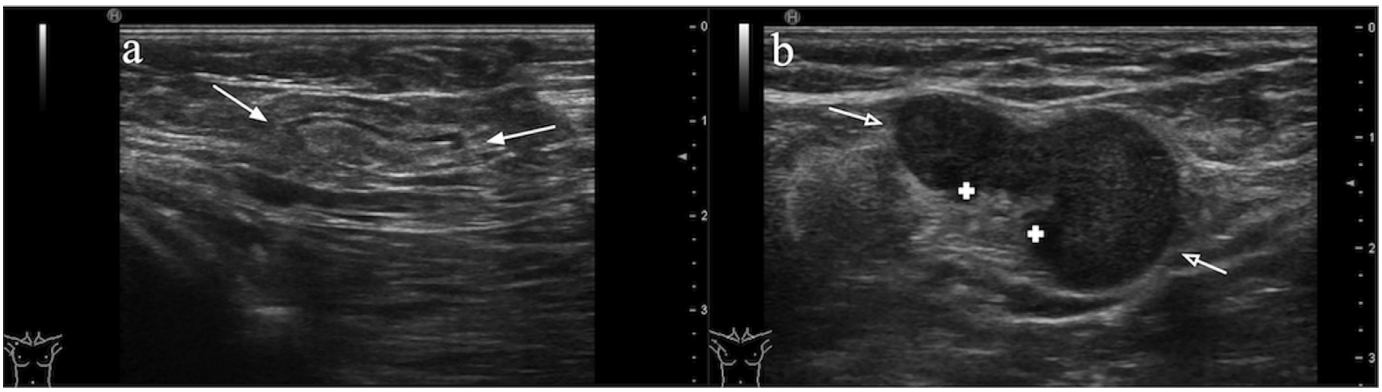

Fig. 1. Illustration of axillary lymph nodes on ultrasound

a) Illustration of unsuspicious ALN on ultrasound. Ultrasound image of a morphologically unsuspicious ALN is oval or round in shape, with a uniform concentric hypoechoic cortex (<3 mm thick), smooth margin and hyperechoic fatty hilum (white arrows).

b) Illustration of suspicious ALN on ultrasound. Ultrasound image of a metastatic ALN is irregular in shape, with a focally thickened hypoechoic cortex and deformity or absence of hyperechoic fatty hilum by compression (white hollow arrows).

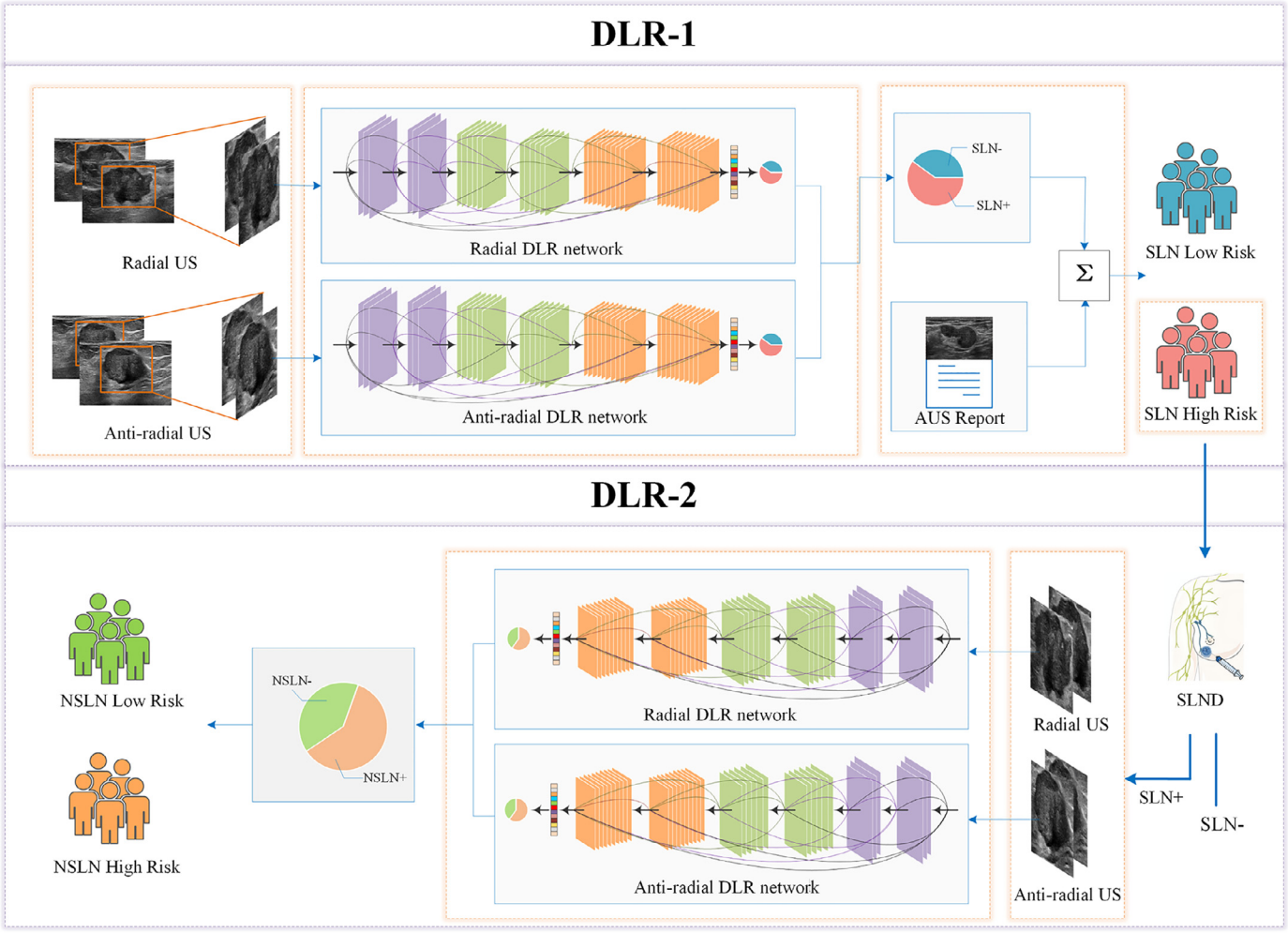

Fig. 2. Workflow of the DLRU

DLRU consists of DLR-1 and DLR-2. The US images at the radial and anti-radial planes were input into the DLR network respectively, which outputs the probability of the SLN metastasis. DLR-1 in the form of the decision tree combined with the DLR model and the AUS report outputs the final predictive risk of the involved SLN. The patient with a high-risk (HR) metastatic SLN will receive an SLND. Finally, the risk of NSLN involvement of the HR SLN patients with positive SLN confirmed by the SLND will be obtained, using DLR-2.

The DLR-2 was also developed using the DLR model to predict the risk of NSLN metastasis. Firstly, two orientations' (620 radial and 641 anti-radial) US images from all SLN+ patients in the training set were used to train the DLR model (details in Table S2, Supplementary material). Secondly, since HR-SLN patients from DLR-1 must undergo subsequent SLND as a treatment in clinical practice. We incorporated the SLND result of these HR-SLN patients into DLR-2 to exclude SLN- patients amongst HR-SLN patients. Therefore, the DLR-2 would output the NSLN metastatic risk of HR-SLN patients with positive SLN (Fig. 2, Fig. S4). Finally, we combined the DLR-1 and DLR-2 as the DLRU, through which some of the SLN+&NSLN- patients could be identified to avoid overtreatment.

The performance of the entire DLRU was validated in an independent test set. All US images from the patients in the test set were employed in this assessment.

Comparison between deep learning radiomics and clinical methods

We investigated the value of clinical features which were previously used as the predictors of the ALN metastasis in improving the performance of DLRU [4]. We selected several tumour-based clinical predictors associated with the SLN metastasis by forward stepwise regression [37]. The regression was applied by using the likelihood ratio test with Akaike's information criterion as the stopping rule. Thereafter, the selected clinical predictors were used for constructing the Clinical model as well as the combination model ($DLR_{+clinical}$) of DLR and clinical features using multivariable logistic regression. We validated the Clinical model and the $DLR_{+clinical}$ model in the test set. Then a comparison of the Clinical model, the DLR model, and the $DLR_{+clinical}$ in detecting the SLN metastasis was made.

Additionally, previously, a few methods were proposed to predict the NSLN metastasis [4,17–20]. Of these methods, the MSKCC nomogram and the Tenon score outperformed others [17,38–40]. Therefore, we summarized the best performance of the MSKCC nomogram and Tenon score in previous studies. To ensure equity, we also randomly selected a subset of the patients with complete clinicopathological data for subgroup analysis of the DLRU and the MSKCC nomogram in predicting the risk of NSLN metastasis (http://www.mskcc.org/nomograms).

Statistical analysis

Descriptive statistics were summarized as mean ± SD or frequencies and percentages. For the quantitative variables, comparisons between groups were made using the Student's *t*-test. Differences of categorical variables were compared using the chi-squared test. The sensitivity, specificity, positive and negative predictive values (PPV, NPV) were calculated to evaluate the performance. The receiver

operating characteristic (ROC) curve was used to show the performance of the model. Statistical significance of differences in sensitivities was assessed by using McNemar's test statistics [41]. A p value $<$ 0.05 was considered statistically significant. All statistical analyses were implemented using SPSS 21, Keras toolkit, R software (version 3.4.1) and Python 3.5.

Results

Prevalence and clinical-pathological characters

Overall, 937 eligible breast cancer patients participated in this study. The prevalence and clinical-pathological characteristics of the patients are described in Table 1. The clinical tumour stage (cT) and biological subtype showed significant differences between the SLN+ and the SLN- patients ($P<$ 0.05 for both; chi-squared test).

The distribution of the detailed ALN status (SLN-, SLN+&NSLN-, and SLN+&NSLN+) is shown in Table S1. Of the SLN+ patients (n = 365), 58% (211/365) were NSLN+ and 42% (154/365) were NSLN- patients (Table S1). The proportion of SLN and NSLN status is consistent with that described in previous studies [4–7,19,42].

Performance of dlru in identifying the metastatic risk of the SLN

As shown in Fig. 2, firstly, the DLRU predicts the risk of SLN involvement. The goal of our study is to ensure reducing overtreatment without adverse impacts on survival. Therefore, high sensitivity is indispensable to ensure essential surgery for positive patients. To better appreciate the sensitivity, we investigated the performance of the clinical characteristics, AUS and the DLR features in diagnosing the risk for SLN invasion. Tumour-based clinical predictors, the cT and biological subtype showed significant correlation to SLN status (P $<$ 0.05, in both sets) and were selected to construct the Clinical model and the combination model ($DLR_{+clinical}$). These two models were compared to the DLR model in detecting the SLN metastasis. Fortunately, the DLR model achieved a significantly higher sensitivity of 87.8% (training set), 89.7% (test set) (P $<$ 0.05, in both sets; McNemar's test) (Table 2). Meanwhile, the AUS report also achieved a reasonable performance (AUC = 0.814, sensitivity = 67.2% in the training set; AUC = 0.825, sensitivity = 70.3% in the test set), which is consistent with previous studies [14,15]. The ROC curve was plotted to demonstrate the comparative results of AUC in Fig. 3a.

Given that the DLR features are derived from the breast tumour and the AUS reflects the information of lymph node, we combined the two into DLRU. Compared with all the above models, the DLRU yields optimal performance with the highest sensitivity (P $<$ 0.05, all models, in both sets; McNemar's test), thereby identifying the most SLN+ patients as HR-SLN (sensitivity = 97.2%, NPV = 97.7%, AUC = 0.88 in training set, sensitivity = 98.4%, NPV = 97% AUC = 0.84 in test set) (Table 2). The sensitivity and specificity of the DLRU are demonstrated in Fig. 4. Ultimately, in the test set (n = 395), of the SLN+ patients (47%, 185/395), 98% (182/185, three patients missed) patients were classified correctly as HR by the DLRU to be assured of a necessary SLND. amongst SLN- patients (n = 210) in the test set, 46% (96/210) patients were correctly classified as LR to avoid an overtreated SLND (Fig. 4).

Performance of the dlru in identifying the metastatic risk of the NSLN

When the DLRU judges the metastatic risk of SLN is HR, it will continue to identify the risk of NSLN metastasis (Fig. 2). In this step, the DLRU yielded remarkable performance with sensitivity = 100%, NPV = 100%, and AUC = 0.91 in the training set. In the test set, the performance dropped slightly with sensitivity = 98.4%, NPV = 91.7% and AUC = 0.81(Table 2). In Fig. 4, the sensitivity and specificity are illustrated. In test set (n = 185), the DLRU stratified 98% (124/126) of the

Table 2
Performance on identifying the risk of SLN and NSLN metastasis.

Set		Method	SEN (%)	SPE (%)	PPV (%)	NPV (%)	AUC	ACC	*P*
SLN	**Trainingset (542)**	DLR	87.8 (82.4–93.1)	60.8 (52.8–68.2)	52.7 (44.7–61.2)	90.9 (84.6–96.1)	0.860 (0.829–0.891)	69.7 (65.7–73.6)	<0.001
		Clinical	3.3 (0.2–7.5)	98.9 (92.5–99.9)	60.0 (54.8–64.2)	67.3 (62.2–73.5)	0.664 (0.622–0.707)	67.2 (63.0–71.1)	<0.001
		$DLR_{+clinical}$	65.6 (52.3–71.8)	89.3 (80.3–95.7)	71.9 (62.5–78.6)	83.6 (78.5–88.8)	0.869 (0.839–0.899)	80.1 (76.5–83.4)	<0.001
		AUS	67.2 (58.9–72.5)	95.6 (90.3–98.4)	88.3 (81.6–91.5)	85.4 (78.2–87.8)	0.814 (0.778–0.850)	86.2 (82.9–88.9)	<0.001
		DLR_{+AUS} (DLR-1)	97.2 (95.7–100)	58.0 (52.5–61.4)	53.5 (47.2–57.3)	97.7 (96.4–100)	0.876 (0.847–0.906)	71.0 (67.0–74.8)	–
	Test set (395)	DLR	89.7 (85.4–93.3)	47.6 (42.7–53.2)	60.1 (55.6–64.4)	84.0 (79.2–89.3)	0.810 (0.768–0.852)	67.3 (62.7–71.9)	<0.001
		Clinical	15.7 (11.3–21.8)	92.9 (86.7–96.2)	65.9 (59.9–70.1)	55.6 (51.0–61.5)	0.674 (0.627–0.721)	56.7 (51.7–61.7)	<0.001
		$DLR_{+clinical}$	69.2 (64.2–75.7)	80.0 (74.6–85.1)	75.3 (70.5–81.2)	74.7 (70.2–81.3)	0.821 (78.0–86.1)	74.9 (70.4–79.1)	<0.001
		AUS	70.3 (65.4–76.9)	94.8 (90.5–98.6)	92.2 (88.2–95.7)	78.4 (72.6–81.8)	0.825 (0.789–0.862)	83.3 (79.2–86.8)	<0.001
		DLR_{+AUS} (DLR-1)	98.4 (96.6–100)	45.7 (41.2–51.8)	61.5 (55.3–67.1)	97.0 (94.2–100)	0.848 (0.811–0.886)	70.4 (65.5–74.8)	–
NSLN	**Training set (175)**	DLR-2 of DLRU	100 (97.4–100)	47.3 (42.0–52.8)	63.6 (58.3–69.5)	100 (98.5–100)	0.909 (0.864–0.953)	72.6 (65.3–79.0)	–
	Test set (182)	DLR-2 of DLRU	98.4 (95.6–99.9)	39.3 (32.5–46.4)	78.5 (71.3–85.6)	91.7 (88.8–97.9)	0.812 (0.740–0.884)	80.2 (73.7–85.7)	–

Abbreviations: DLR, deep learning radiomics model; DLRU, deep learning radiomics of ultrasonography; Clinical, clinical model; $DLR_{+clinical}$, the model combined with DLR and clinical features; AUS, axillary ultrasonography report; DLR_{+AUS}, the model combined with the DLR model and the AUS report; AUC, area under the receiver operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value.
P represents the sensitivity differences between each model and DLR_{+AUS} (McNemar's test).
Data in parentheses are the 95% confidence interval.

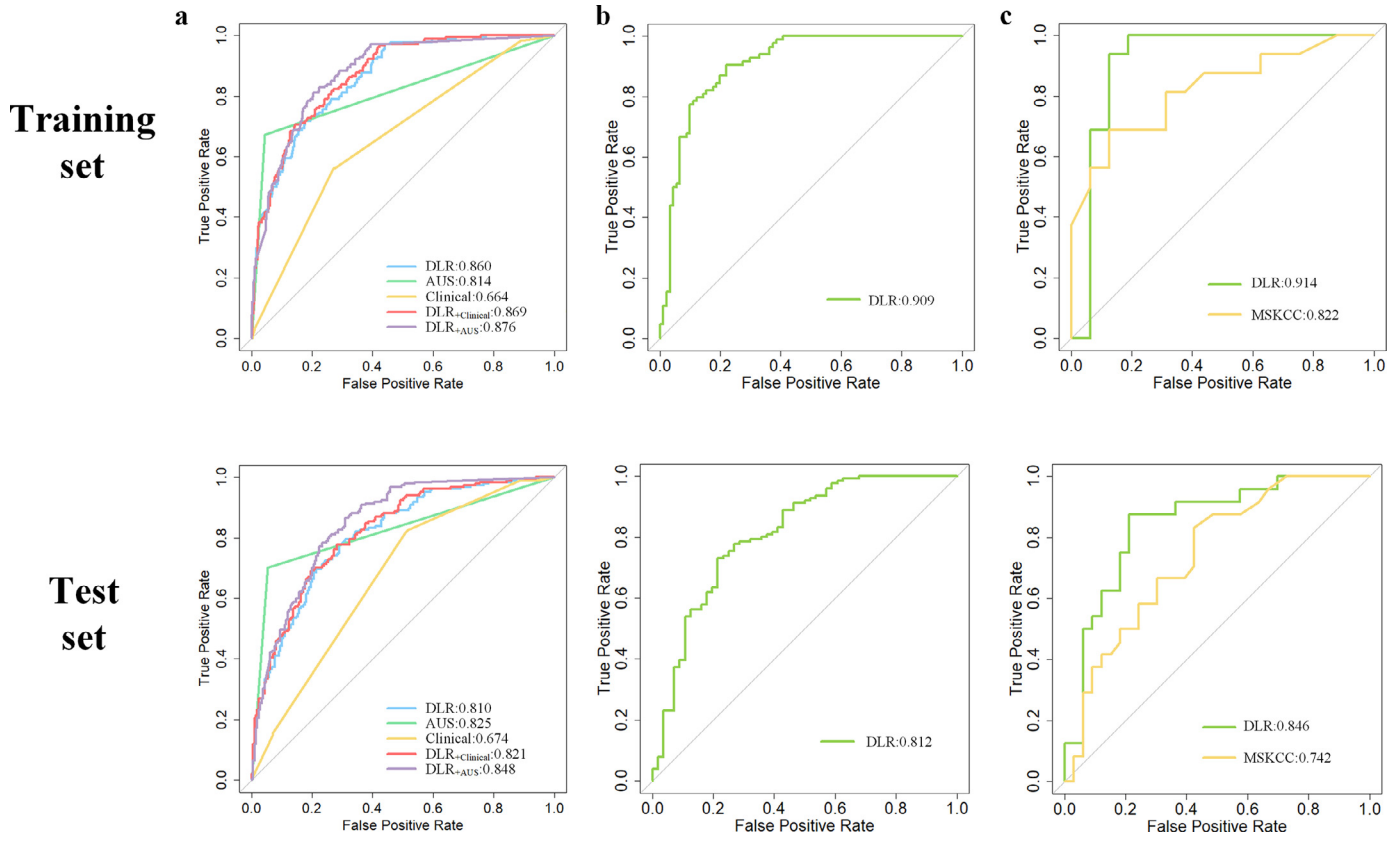

Fig. 3. ROC curves of the models for identifying the metastatic risk of SLN or NSLN in the training and test sets

a) ROC curves of different models for predicting the SLN status.

b) ROC curves of the Submodel-2 of DLRU for predicting the risk of NSLN metastasis.

c) ROC curves of subgroup analysis between the DLRU and the MSKCC nomogram for evaluating the risk of NSLN metastasis.

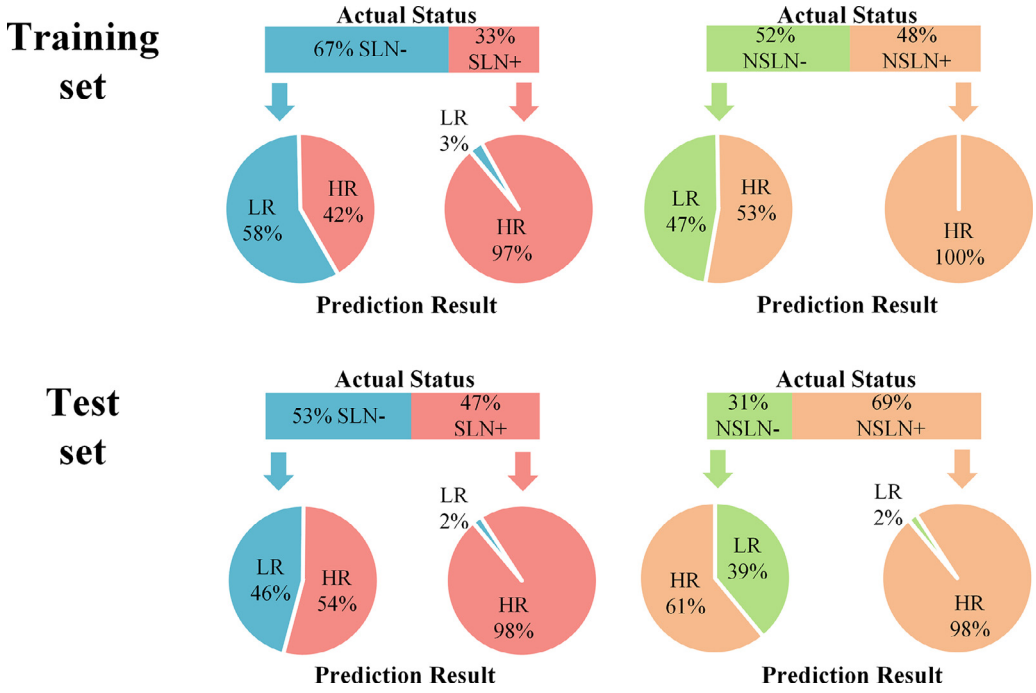

Fig. 4. Performance of the DLRU in the stepwise evaluation of the risk of SLN and NSLN involvement in training and test sets

The upper panel shows the actual status distribution percentage of SLN and NSLN, and the lower panel shows the prediction results outputted by the DLRU in the training and the test sets. For example, in the training set, 67% of the patients were SLN- and 33% were SLN+. The DLRU classifies 58% of SLN- patients as low-risk (LR) and 97% of SLN+ patients as high-risk (HR).

SLN+&NSLN+ patients (n = 128, two patients missed by DLR-1) into a HR-SLN&HR-NSLN category. A necessary ALND could be ensured for this category. Moreover, in the test set, of the SLN+ patients (n = 185) without tumour-bearing in the NSLN (31%, 57/185), 39% (22/56, one patient missed by DLR-1) cases were categorized as HR-SLN&LR-NSLN by DLRU for whom the morbidities of over-treated ALND could be avoided (Fig. 4). The ROC curve of DLR-2 is shown in Fig. 3b.

We also compared DLRU with the MSKCC nomogram and Tenon score in Table 3 [17,38–40]. In terms of sensitivity, PPV, NPV and AUC, the DLRU outperforms the best performance of these two models in previous studies. Moreover, in the subgroup analysis, the sensitivity of the DLRU (1.00, 0.914 in the training set; 0.958, 0.846 in the test set) was consistently higher than that of the MSKCC nomogram (0.25, 0.82 in the training set; 0.42, 0.74 in the test set) ($P < 0.05$, in both sets; McNemar's test) (Table 3, Fig. 3c). Note that the sensitivity of DLRU for identifying HR-NSLN was higher than the MSKCC nomogram, though the specificity of DLRU is lower. The reason is that we deliberately designed the DLRU to minimize the chance of missing the high-risk axillary invasion to ensure the treatment, that is, to achieve high sensitivity rather than specificity.

To better comprehend the DLRU, we visualize the response areas to DLR model in predicting the status of NSLN by applying the Gradient-weighted Class Activation Mapping (Fig. 5a, Fig. S5). We found that in US images of different NSLN statuses, the location of the strong response area obtained from the DLR model may change. In most US images of the NSLN- patients, the strong response areas often tend to cluster on the edge of the tumour. And in most US images of the NSLN+ patients, the strong response areas usually cluster in the tumour. To some tend, this may explain the discrimination ability of the DLR, which is consistent with the previous study [29].

Comparing the benefit of DLRU with the current clinical practice

Current clinical practice is based on SLN assessment for all patients by SLN biopsy or SLND. When metastatic SLN is detected, a subsequent ALND will be performed. These would lead to SLN- or SLN+&NSLN- patients undergone overtreated SLND or ALND. However, by two-step outputs of DLRU, the patient could be assigned to one of the three groups corresponding to three ideal axillary management groups: LR-SLN, HR-SLN&LR-NSLN, and HR-SLN&HR-NSLN (see Materials and Methods) (Fig. S4). In this way, using DLRU, some of SLN- or SLN+&NSLN- patients may avoid overtreatment.

As shown in the left panel of Fig. 5b, in the whole study cohort, when we made an operative proposal according to the current clinical management, it resulted in 77.4% (61% SLN- and 16.4% SLN+&NSLN-) of the patients being over-treated with unnecessary SLND or ALND and only 22.6% NSLN+ of the patients received a suitable ALND. In contrast, when we analysed these same patients using the DLRU, it stratified 32.7% of the patients correctly into the LR-SLN group, thus avoiding overtreatment of SLND (Fig. 5b, right panel). The DLRU also stratified 42% (6.9%/16.4%) of the SLN+&NSLN- patients into HR-SLN&LR-NSLN to receive suitable surgery (SLND) and omit overtreatment (ALND). Furthermore, almost all (98.2%, 22.2%/22.6%) of the SLN+&NSLN+ patients were classified as HR-SLN&HR-NSLN to ensure necessary ALND. Eventually, only 37.1% of all the breast cancer patients were mistakenly stratified into HR, leading to overtreatment. The application of DLRU, therefore, may have the potential to further reduce morbidity and economic costs associated with current axillary management (37.1% vs. 77.4%; Fig. 5b),

Discussion

The NCCN Guidelines recommend that breast cancer patient with positive SLN can undergo SLND instead of ALND if the metastatic risk of NSLN is low [8]. However, whether SLN or NSLN, the status can only be obtained by the invasive SLND or ALND, They both have

Table 3
Comparison and subgroup analysis of models in identifying NSLN metastasis [17,40].

Method	Set	SEN (%)	SPE (%)	PPV (%)	NPV (%)	AUC	ACC	*P*
MSKCC nomogram	previous research [17]	93.1	–	–	–	0.78	–	–
Tenon score	previous research [17,40]	92.1	70.1	54.7	95.8	0.81	–	–
DLR model proposed	Training set	100 (97.4–100)	47.3 (42.0–52.8)	63.6 (58.3–69.5)	100 (98.5–100)	0.909 (0.864–0.953)	72.6 (65.3–79.0)	–
	Test set	98.4 (95.6–99.9)	39.3 (32.5–46.4)	78.5 (71.3–85.6)	91.7 (88.8–97.9)	0.812 (0.740–0.884)	80.2 (73.7–85.7)	–
DLR model proposed	Trainingset (32)	100 (91.2–100)	50.0 (39.2–61.8)	66.7 (50.3–75.0)	100 (88.5–100)	0.914 (0.787–1)	75.0 (56.7–88.5)	<0.001
MSKCC nomogram		25.0 (12.4–35.8)	100 (95.5–100)	100 (96.1–100)	57.1 (50.1–64.2)	0.822 (0.675–0.970)	62.5 (43.7–78.9)	
DLR model proposed	Test set (57)	95.8 (91.1–99.3)	33.3 (27.8–39.5)	51.1 (42.8–57.6)	91.7 (87.2–97.8)	0.846 (0.742–0.950)	59.7 (45.8–72.4)	<0.001
MSKCC nomogram		41.7 (35.9–46.4)	84.9 (79.9–91.5)	66.7 (59.1–72.8)	66.7 (61.0–72.3)	0.742 (0.614–0.870)	66.7 (52.9–78.6)	

Abbreviations: MSKCC, Memorial Sloan-Kettering Cancer Centre; Tenon score, a scoring system; AUC, area under the receiver operating characteristic curve; ACC, accuracy; SEN, sensitivity; SPE, specificity; PPV, positive predictive value; NPV, negative predictive value; DLR model, deep learning radiomics model;.
P represents the sensitivity difference between MSKCC nomogram and DLR model (McNemar's test).
Data in parentheses are the 95% confidence interval.

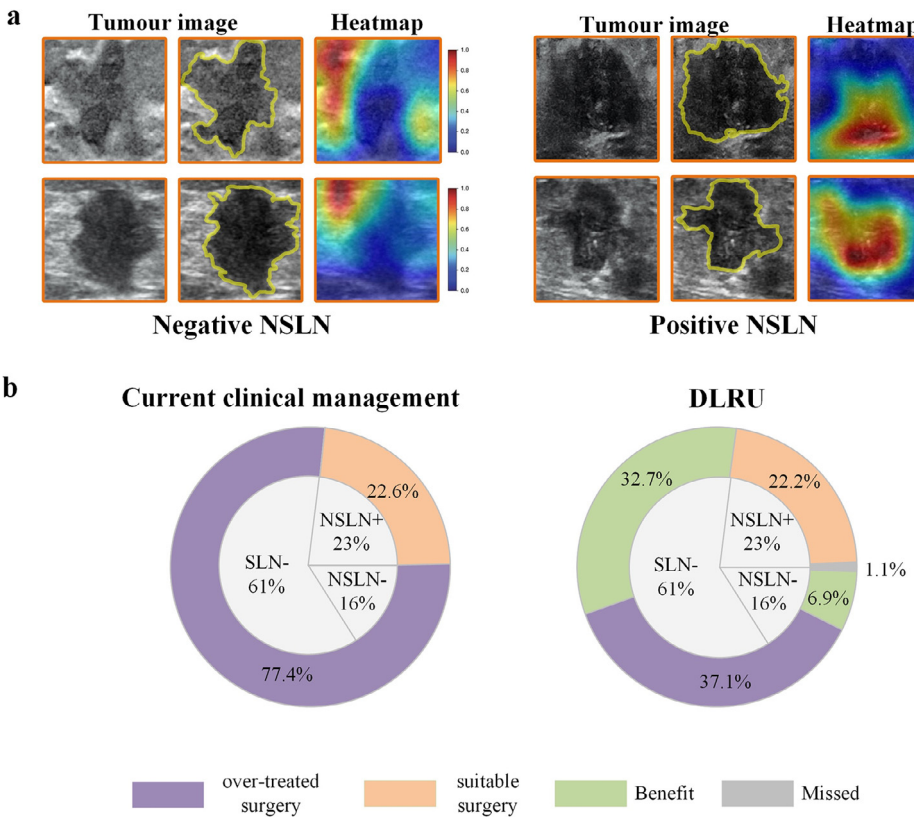

Fig. 5. DLRU analysis

a) Visualization of the DLRU

The strong response area (red area) obtained from the DLR model in US images of different NSLN statuses.

b) Overall benefit by DLRU

These two ring diagrams respectively show the proportion of patients in the study who will receive different treatments according to DLRU or current clinical management. The inner ring represents the actual distribution percentage of ALN status of all patients in this study. The outer ring of the left ring diagram shows the treatment based on the current clinical management, leading to the overtreatment of 77•4% patients, and suitable treatment of 22•6% patients. The outer ring of the right ring diagram shows axillary surgery decisions made by the DLRU, 37•1% of the patients will undergo overtreatment, and 62•6% of all patients (SLN- or SLN+&NSLN-) will benefit and receive suitable surgery.

complications and morbidities in varying degrees [2,3,11,43]. In this study, a novel multicentre model based on DLR and conventional two-dimensional US imaging, called DLRU, was developed for a personalized evaluation of the risk of SLN and NSLN invasions. The most general US imaging was utilized in our study to simplify the service conditions of the DLRU. By this means, the DLRU could probably be helpful even in the basic hospitals or the medically underdeveloped areas.

In the clinic, an essential criterion can be helpful to measure the utility of a method: false-negative rate (FNR, 1- sensitivity). An FNR with 5% is widely used as a target value when previous studies assessing a model, due to the FNR of ALND is close to 5% [17,44]. Besides, the FNR of the SLND in the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-32 and other randomized trials were less than 9.8% [7,45]. So when applying the model, the target FNR adjusted to less than 10% has been considered clinically acceptable [7,17,44,45]. In both sets, the DLRU yields sufficient low FNR in assessing the risk of the SLN and NSLN involvement, which is less than 5%. This indicates a very low likelihood of missed breast cancer with ALN metastasis by the model, thus ensuring the essential surgery for these patients. And previous studies demonstrated that SLND for patients with limited SLN metastasis can achieve comparable survival outcome with the ALND and less morbidities, such as wound infection, seroma, paraesthesia and lymphedema, which translates into a better quality of life (QoL) [2,3,6,7,9–11,43]. Therefore, with a sufficiently low FNR, HR-SLN&LR-NSLN patients identified by DLRU may be able to safely convert ALND to SLND, thereby obtaining comparable survival outcomes and better QoL.

We have to mention that though FNR was consistent with the SLND, a small portion of breast cancer patients with ALN metastasis that were misclassified into LR group were non-negligible. Similarly, in the current clinical treatment, there are also concerns about the missed diagnosis of SLND. To ensure the treatment for some patients suspected of missed diagnosis, the NCCN panel notes that postoperative axillary radiotherapy (ART) can be performed on these patients for regional control of the disease [8]. This recommendation can also be applied to patients using DLRU. Moreover, ART has been proven to provide axillary control comparable to ALND for some patients and result in better QoL [42]. By performing ART, a small but non-negligible portion of breast cancer patients who were misjudged by the model may ensure the equivalent locoregional control.

We investigated several prediction models to examine the possibility of improving sensitivity (Table 2). The Clinical model, even the $DLR_{+clinical}$ model is inferior to the DLR model in identifying patients with positive SLN. Meanwhile, either the DLR model or the AUS report achieved a significantly lower sensitivity than DLR-1 composed by the two. We assume that such results are probably caused by the following reasons: a) Tumour-based clinical features are the external manifestations of the intra-tumour heterogeneity (ITH), which is a common feature in solid tumour cancers [46]. Multiple studies have demonstrated that DLR features acquired from macroscopic medical images can comprehensively reflect the innate ITH information of the tumour [25,47]. Therefore, the tumour-based clinical features may be redundant in improving the sensitivity of the DLR model. b). The AUS report contains general morphology information of the ALN, which is independent of the primary tumour and

complements the axillary features. Therefore, the performance for predicting the metastatic SLN is well improved by incorporating the AUS report as a factor into DLR-1.

Considering the information in previous studies, our study may be the first attempt to identify NSLN metastasis, using the DLR features derived from US imaging. Many studies constructed models to assess the NSLN status using SLN pathological indicators such as the number of positive and negative SLNs, the detective method of the SLN, the size of the largest SLN metastasis, and the presence of extra-capsular extension of the SLN metastasis [4,18–20]. We compared the performance of DLRU to the MSKCC nomogram and the Tenon score which were demonstrated superior to other methods [17,38,39]. It was found that the DLRU performed invariably better in identifying positive patients (Table 3). The DLR shows remarkable performance in this task of extracting information connected to NSLN metastasis from the breast cancer US imaging, as in other studies [25,34].

With a rigid pathological examination of the SLN, the ACOSOG Z0011and IBCSG 23–01 trials suggest that if patients meet a series of criteria simultaneously, omitting further ALND after tumour-positive SLND does not decrease the survival rate [6,9,10,12]. Nevertheless, the accuracy of the pathological assessment varied between hospitals. In less developing countries, a large population base or resource limitations in terms of technological capacity and personnel, leading to the pathological examination cannot reach such strict standards. Moreover, all the patients enrolled in these trials underwent breast-conserving surgery followed by whole-breast radiation. According to the preliminary data released at the 15th Chinese breast cancer conference, only 22% of the breast cancer patients in China have undergone breast-conserving surgery; this statistic is similar to that obtained in other countries [48,49]. These limit the generalization of the conclusion obtained from the trials.

To make our model generalizable, we have tried to intentionally minimize our exclusion criteria. All patients in our study suffered from primary breast cancer, and there was no restriction on tumour staging, pathological type and molecular subtype, so as to include more diverse breast cancers. This may make the heterogeneous population of patients with primary breast cancer close to the actual clinical reality. And we used the most accessible and universal medical imaging method, US, making the model helpful even in medical resource-limited areas. To ensure that the overtreatment of negative patients is appropriately reduced without adversely affecting the necessary surgery for positive patients, the DLRU gets a reasonably stable low FNR in two completely different hospitals. This may demonstrate the generalization capability of the DLRU. Additionally, nowadays, since the individualized treatment options of the axilla are under research, patient-specific prediction of axillary disease is getting more attention. The ACOSOG Z0011 trial, the IBCSG 23–01 trial and the AMAROS trial implicate that the role of intraoperative SLN pathology analysis in the clinical context is diminishing [4,6,9,10]. In this perspective, the DLR might be able to promote the trend of simplifying axillary surgery uniquely.

Despite the outstanding performance of the DLRU, this study still has a few limitations in terms of the retrospective study, the limited number of hospitals, the further optimizable DLRU model and the inherent defects of the US. Firstly, since this is a retrospective study, excluding patients with incomplete data required for the study is inevitable and may lead to bias. Therefore, prospective studies are needed in future studies to confirm the realistic performance of the model and overcome the potential bias derived from retrospective data. Also, subsequent studies can involve more hospitals and more diversified breast cancers to create a real-world distribution of patients to better train the DLRU for better stability. Secondly, multi-modal US such as Colour Doppler US, ultrasonic elastography and the contrast-enhanced US reflect varied intrinsic information of the tumour. In further research, we will focus on improving the performance of DLRU by involving more patients and multi-modal US images to maintain high sensitivity and increase specificity. Thirdly, although we adopted the relatively simple method that analysed a single image at a time to make our model better adapt to the data collection preferences of different hospitals and doctors, this prevents us from capturing the correlation between images. Therefore, in subsequent research, we will further explore better methods to analyse multiple consecutive images at the same time. Finally, due to the inherent limitations of ultrasound, our research inevitably lacks a small number of breast tumours that are invisible in the ultrasound. Therefore, further studies involving more medical imaging methods remains to be researched.

In conclusion, we present a novel DLRU model for the identification of the metastatic risk in SLN and NSLN in primary breast cancer. The DLRU demonstrates good performance, thus, implying the promising potential of the DLR in risk-assessment of the ALN. If possible, pre-surgical use of the DLRU may lead to a reduction in morbidities of ALND or SLND without adverse impact on survival. However, further studies on a larger data set with more centres are desirable.

Data sharing

Due to the privacy of patients, the data related to patients cannot be available for public access but can be obtained from the corresponding author on reasonable request approved by the institutional review board of Harbin Medical University. (JW.T. jwtian2004@163.com)

Acknowledgements

We would like to thank Dr. Wen Cheng from the Department of Medical Ultrasonic at the third affiliated hospital of HMU for enabling and facilitating patient recruitment and contributing data to the study.

Funding Sources

This study was supported by the National Natural Science Foundation of China (Grant Nos. 81974265, 81701705, 81630048, 81271647 and 81901761), National Outstanding Youth Science Fund Project of National Natural Science Foundation of China (Grant No. 81101103), Heilongjiang Provincial Postdoctoral Science Foundation (Grant No. LBH-Z17174), Scientific research project of Heilongjiang Health Committee (Grant No. 2019–050), Postgraduate Research &Practice Innovation Program of Harbin Medical University (Grant No. YJSSJCX2019–08HYD). The funders had no role in study design, data collection, data analysis, interpretation, or writing of the report.

Declaration of Interests

The authors declared no conflict of interest.

Author Contributions

Conceptualization: Xu Guo, Zhenyu Liu, Lei Zhang, Ying Wang, Jie Tian and Jiawei Tian;

Investigation and Data curation: Xu Guo, Lei Zhang, Ying Wang, Ziyao Li, Jiaxin Shi, Jing Zhang, Hao Cui, Tong Wu and Jiawei Tian;

Methodology, Formal analysis and Validation: Zhenyu Liu, Caixia Sun and Xu Guo;

Writing and editing: Xu Guo, Zhenyu Liu, Caixia Sun, Jie Tian and Jiawei Tian.

Resources and Supervision: Jie Tian and Jiawei Tian.

All the authors final reviewed the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2020.103018.

References

[1] Braithwaite LR. The flow of lymph from the ileocæcal angle, and its possible bearing on the cause of duodenal and gastric ulcer. Br J Surg 1923;11:7–26.
[2] Purushotham AD, Upponi S, Klevesath MB, et al. Morbidity After Sentinel Lymph Node Biopsy in Primary Breast Cancer: results From a Randomized Controlled Trial. J Clin Oncol 2005;23:4312–21.
[3] Lucci A, McCall LM, Beitsch PD, et al. Surgical complications associated with sentinel lymph node dissection (SLND) plus axillary lymph node dissection compared with SLND alone in the American College of Surgeons Oncology Group trial Z0011. J Clin Oncol 2007;25:3657–63.
[4] Meretoja TJ, Leidenius MHK, Heikkilä PS, et al. International multicenter tool to predict the risk of nonsentinel node metastases in breast cancer. J Natl Cancer Inst 2012;104:1888–96.
[5] Veronesi U, Paganelli G, Galimberti V, et al. Sentinel-node biopsy to avoid axillary dissection in breast cancer with clinically negative lymph-nodes. Lancet 1997;349:1864–7.
[6] Galimberti V, Cole BF, Viale G, et al. Axillary dissection versus no axillary dissection in patients with breast cancer and sentinel-node micrometastases (IBCSG 23-01): 10-year follow-up of a randomised, controlled phase 3 trial. Lancet Oncol 2018;19:1385–93.
[7] Veronesi U, Paganelli G, Viale G, et al. A Randomized Comparison of Sentinel-Node Biopsy with Routine Axillary Dissection in Breast Cancer. N Engl J Med 2003;349:546–53.
[8] National Comprehensive Cancer Network. (NCCN) Clinical practice guidelines in oncology. Breast Cancer Version3.2019. *Natl Compr Cancer Network Clin Pract Guidel Oncol (accessed Sept 6, 2019)*https://www.nccn.org/professionals/physician_gls/default.aspx#breast.
[9] Giuliano AE, Ballman K V, McCall L, et al. Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: the ACOSOG Z0011 (Alliance) randomized clinical trial. JAMA - J Am Med Assoc 2017;318:918–26.
[10] Giuliano AE, Ballman K, McCall L, et al. Locoregional Recurrence After Sentinel Lymph Node Dissection With or Without Axillary Dissection in Patients With Sentinel Lymph Node Metastases. Ann Surg 2016;264:413–20.
[11] Rao R, Euhus D, Mayo HG, Balch C. Axillary Node Interventions in Breast Cancer. JAMA 2013;310:1385.
[12] Galimberti V, Cole BF, Zurrida S, et al. Axillary dissection versus no axillary dissection in patients with sentinel-node micrometastases (IBCSG 23–01): a phase 3 randomised controlled trial. Lancet Oncol 2013;14:297–305.
[13] DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. CA Cancer J Clin 2019;69:438–51.
[14] Tucker NS, Cyr AE, Ademuyiwa FO, et al. Axillary ultrasound accurately excludes clinically significant lymph node disease in patients with early stage breast cancer. Ann Surg 2016;264:1098–102.
[15] Cools-Lartigue J, Meterissian S. Accuracy of axillary ultrasound in the diagnosis of nodal metastasis in invasive breast cancer: a review. World J Surg 2012;36:46–54.
[16] Kim GR, Choi JS, Han B-K, et al. Preoperative Axillary US in Early-Stage Breast Cancer: potential to Prevent Unnecessary Axillary Lymph Node Dissection. Radiology 2018;288:55–63.
[17] Coutant C, Olivier C, Lambaudie E, et al. Comparison of Models to Predict Nonsentinel Lymph Node Status in Breast Cancer Patients With Metastatic Sentinel Lymph Nodes: a Prospective Multicenter Study. J Clin Oncol 2009;27:2800–8.
[18] Van Zee KJ, Manasseh DME, Bevilacqua JLB, et al. A nomogram for predicting the likelihood of additional nodal metastases in breast cancer patients with a positive sentinel node biopsy. Ann Surg Oncol 2003;10:1140–51.
[19] Viale G, Maiorano E, Pruneri G, et al. Predicting the risk for additional axillary metastases in patients with breast carcinoma and positive sentinel lymph node biopsy. Ann Surg 2005;241:319–25.
[20] Degnim AC, Reynolds C, Pantvaidya G, et al. Nonsentinel node metastasis in breast cancer patients: assessment of an existing and a new predictive nomogram. Am J Surg 2005;190:543–50.
[21] Pal A, Provenzano E, Duffy SW, Pinder SE, Pumshotham AD. A model for predicting non-sentinel lymph node metastatic disease when the sentinel lymph node is positive. Br J Surg 2008;95:302–9.
[22] Han L, Zhu Y, Liu Z, et al. Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer. Eur Radiol 2019;29:3820–9.
[23] Liu C, Ding J, Spuhler K, et al. Preoperative prediction of sentinel lymph node metastasis in breast cancer by radiomic signatures from dynamic contrast-enhanced MRI. J Magn Reson Imaging 2019;49:131–40.
[24] Golden JA. Deep learning algorithms for detection of lymph node metastases from breast cancer helping artificial intelligence be seen. JAMA - J Am Med Assoc 2017;318:2184–6.
[25] Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. CA Cancer J Clin 2019;0:1–31.
[26] Hosny A, Aerts HJWL. Artificial Intelligence for Global Good. Science 2019;366:955–6.
[27] Gierach GL, Li H, Loud JT, et al. Relationships between computer-extracted mammographic texture pattern features and BRCA1/2mutation status: a cross-sectional study. Breast Cancer Res 2014;16:424.
[28] Zhou L, Wu X-L, Huang S-Y, et al. Lymph Node Metastasis Prediction from Primary Breast Cancer US Images Using Deep Learning. Radiology 2020;294:19–28.
[29] Zheng X, Yao Z, Huang Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. Nat Commun 2020;11:1236.
[30] Guidelines AP, Ultrasound B. AIUM Practice Guideline for the Performance of a Breast Ultrasound Examination. J Ultrasound Med 2009;28:105–9.
[31] Mendelson EB, Böhm-Vélez M, Berg WA, et al. ACR BI-RADS® Ultrasound. ACR BI-RADS® atlas, breast imaging reporting and data system. reston, va. American College of Radiology; 2013.
[32] Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely Connected Convolutional Networks. In: 2017 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, 2017: 2261–9.
[33] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8.
[34] Wang S, Shi J, Ye Z, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. Eur Respir J 2019;53. doi: 10.1183/13993003.00986-2018.
[35] Berg JW. The significance of axillary node levels in the study of breast carcinoma. Cancer 1955;8:776–8.
[36] Veronesi U, Paganelli G, Viale G, et al. Sentinel Lymph Node Biopsy and Axillary Dissection in Breast Cancer: results in a Large Series. JNCI J Natl Cancer Inst 1999;91:368–73.
[37] Sauerbrei W, Boulesteix AL, Binder H. Stability investigations of multivariable regression models derived from low-and high-dimensional data. J Biopharm Stat 2011;21:1206–31.
[38] Van Den Hoven I, Kuijt G, Roumen R, Voogd A, Steyerberg EW, Vergouwe Y. A head to head comparison of nine tools predicting non-sentinel lymph node status in sentinel node positive breast cancer women. J Surg Oncol 2015;112:133–8.
[39] Rouzier R, Uzan C, Rousseau A, et al. Multicenter prospective evaluation of the reliability of the combined use of two models to predict non-sentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: the MSKCC nomogram and the Tenon score. Results of the NOTEGS. Br J Cancer 2017;116:1135–40.
[40] Coutant C, Rouzier R, Fondrinier E, et al. Validation of the Tenon breast cancer score for predicting non-sentinel lymph node status in breast cancer patients with sentinel lymph node metastasis: a prospective multicenter study. Breast Cancer Res Treat 2009;113:537–43.
[41] Agresti A. An introduction to categorical data analysis: second edition. Gainesville, Florida: Wiley; 2006. doi: 10.1002/0470114754.
[42] Donker M, van Tienhoven G, Straver ME, et al. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): a randomised, multicentre, open-label, phase 3 non-inferiority trial. Lancet Oncol 2014;15:1303–10.
[43] Land SR, Kopec JA, Julian TB, et al. Patient-reported outcomes in sentinel node-negative adjuvant breast cancer patients receiving sentinel-node biopsy or axillary dissection: national Surgical Adjuvant Breast and Bowel Project phase III protocol B-32. J Clin Oncol 2010;28:3929–36.
[44] Werkoff G, Lambaudie E, Fondrinier E, et al. Prospective Multicenter Comparison of Models to Predict Four or More Involved Axillary Lymph Nodes in Patients With Breast Cancer With One to Three Metastatic Sentinel Lymph Nodes. J Clin Onc. 2009;27:5707–12.
[45] Krag DN, Anderson SJ, Julian TB, et al. Technical outcomes of sentinel-lymph-node resection and conventional axillary-lymph-node dissection in patients with clinically node-negative breast cancer: results from the NSABP B-32 randomised phase III trial. Lancet Oncol 2007;8:881–8.
[46] McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: past, Present, and the Future. Cell 2017;168:613–28.
[47] Liu Z, Li Z, Qu J, et al. Radiomics of multiparametric MRI for pretreatment prediction of pathologic complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study. Clin Cancer Res 2019;25:3538–47.
[48] Wong WJ, Mosiun JA, Hidayati Z, et al. Low Breast Conserving Surgery (BCS) rates in public hospitals in Malaysia: the effect of stage and ethnicity. Breast 2019;46:136–43.
[49] Trocchi P, Kuss O, Kääb-Sanyal V, Heidinger O, Stang A. Trends in surgical treatment for breast cancer in Germany after the implementation of the mammography screening program. Eur J Epidemiol 2019;34:1143–50.