



Application of deep learning to predict underestimation in ductal carcinoma in situ of the breast with ultrasound

Lang Qian^{1,2#}, Zhikun Lv^{3,4#}, Kai Zhang^{1,2#}, Kun Wang^{3,4}, Qian Zhu^{1,2}, Shichong Zhou^{1,2}, Cai Chang^{1,2}, Jie Tian^{3,4,5}

¹Department of Ultrasonography, Fudan University Shanghai Cancer Center, Shanghai, China; ²Department of Oncology, Fudan University, Shanghai Medical College, Shanghai, China; ³CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China; ⁴School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China; ⁵Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing, China

Contributions: (I) Conception and design: J Tian, S Zhou, L Qian, Z Lv; (II) Administrative support: J Tian, C Chang, S Zhou, K Wang; (III) Provision of study materials or patients: C Chang, S Zhou; (IV) Collection and assembly of data: L Qian, K Zhang, Q Zhu; (V) Data analysis and interpretation: Z Lv, L Qian, K Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Jie Tian. CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China. Email: tian@ieec.org; Shichong Zhou. Department of Ultrasonography, Fudan University Shanghai Cancer Center, Shanghai 200032, China. Email: sczhou@fudan.edu.cn.

Background: To develop an ultrasound-based deep learning model to predict postoperative upgrading of pure ductal carcinoma in situ (DCIS) diagnosed by core needle biopsy (CNB) before surgery.

Methods: Of the 360 patients with DCIS diagnosed by CNB and identified retrospectively, 180 had lesions upstaged to ductal carcinoma in situ with microinvasion (DCISM) or invasive ductal carcinoma (IDC) postoperatively. Ultrasound images obtained from the hospital database were divided into a training set (n=240) and validation set (n=120), with a ratio of 2:1 in chronological order. Four deep learning models, based on the ResNet and VggNet structures, were established to classify the ultrasound images into postoperative upgrade and pure DCIS. We obtained the area under the receiver operating characteristic curve (AUROC), specificity, sensitivity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) to estimate the performance of the predictive models. The robustness of the models was evaluated by a 3-fold cross-validation.

Results: Clinical features were not significantly different between the training set and the test set (P value >0.05). The area under the receiver operating characteristic curve of our models ranged from 0.724 to 0.804. The sensitivity, specificity, and accuracy of the optimal model were 0.733, 0.750, and 0.742, respectively. The three-fold cross-validation results showed that the model was very robust.

Conclusions: The ultrasound-based deep learning prediction model is effective in predicting DCIS that will be upgraded postoperatively.

Keywords: Artificial intelligence (AI); ductal carcinoma in situ (DCIS); core needle biopsy (CNB); prediction of upstaging

Submitted May 15, 2020. Accepted for publication Nov 12, 2020.

doi: 10.21037/atm-20-3981

View this article at: <http://dx.doi.org/10.21037/atm-20-3981>

Introduction

In recent years, with the global spread and development of breast cancer screening, the detection rate of ductal carcinoma *in situ* (DCIS) is increasing, accounting for about 20% of diagnosed breast cancers (1). Presently, the main treatment of DCIS is surgery. For a breast mass, the treatment would include mastectomy or lumpectomy plus radiation therapy. According to the American Society of Clinical Oncology guidelines, patients diagnosed with pure DCIS by core needle biopsy (CNB) before surgery should undergo sentinel lymph node biopsy (SLNB) if they choose mastectomy (2); patients who undergo lumpectomy should undergo SLNB if DCIS is upgraded postoperatively.

The treatment of DCIS such as the management of lymph nodes is controversial. Theoretically, pure DCIS does not have axillary lymph node metastases. However, approximately 12–32% of cases diagnosed by CNB before surgery involve upstaging to a microinvasion; this indicates the invasion of cancer cells beyond the basement membrane into at least 1 mm of the adjacent tissue or the diagnosis of an invasive cancer on postoperative specimen analysis (3–6). The major cause of upstaging cannot be determined by CNB, and imaging manifestations, biopsy techniques, and the DCIS size could also affect the preoperative diagnosis (3–6). Therefore, overtreatment and undertreatment may occur in the management of axillary lymph nodes in patients with DCIS. For example, clinicians may perform SLNB at the time of initial surgery if DCIS was upstaged before surgery. On the other hand, patients who choose lumpectomy may require a second operation for SNLB, which may increase the financial and psychological burden.

To prevent overtreatment caused by over-diagnosis of DCIS in clinical practice, prospective studies on whether patients can be treated with active monitoring, follow-up, radiotherapy, and other non-surgical treatments instead of traditional surgical treatments are currently ongoing in the United Kingdom [Low Risk DCIS trial (LORIS)] (7) and in the United States [the comparison of operative versus medical endocrine therapy for low-risk DCIS trials (COMET)] (8). DCIS at high risk of stromal invasion should be excluded before non-surgical treatment is considered. Therefore, predictors of postsurgical upstaging of preoperatively diagnosed pure DCIS by CNB are critical.

Presently, DCIS is mainly screened using mammography. Its main imaging manifestation is the presence of clustered microcalcifications, but this feature is not unique to DCIS. Therefore, it is difficult to distinguish DCIS from invasive

carcinoma using imaging studies. In addition, postoperative upstaging is observed in patients diagnosed with DCIS using puncture biopsy. Distinction between DCIS and invasive carcinoma before surgery has been addressed previously. Recently, research has focused on corresponding clinical factors as predictors of postoperative upstaging of pure DCIS diagnosed preoperatively by CNB (3,9,10). However, the evaluation of some of these clinical factors have been subjective, and the factors are difficult to apply in clinical practice. Recently, with the development of artificial intelligence (AI), researchers have evaluated models that can extract effective features through large-scale images and clinical data to predict postoperative upgrading of pure DCIS diagnosed by CNB. Currently, all AI prediction studies are based on mammography or magnetic resonance imaging (MRI). While ultrasonography is commonly used in breast examination, there are no AI studies that use ultrasound images to predict the postoperative upgrade of pure DCIS diagnosed by CNB. The purpose of this study was to predict the postoperative upgrading of pure DCIS, diagnosed by preoperative CNB using deep learning based on two-dimensional ultrasound images. We presented the following article in accordance with the STARD 2015 reporting checklist (available at <http://dx.doi.org/10.21037/atm-20-3981>).

Methods

Patients

For optimal performance of the convolution network model, the sample sizes in the two groups (upstaged and pure DCIS) should be equal (11). Previous research has shown a wide variability of the number of cases in the two groups, with an upgrade rate of pure DCIS after surgery of 12–32% (3–6). Therefore, to balance the number of upstaged and pure DCIS, we retrospectively enrolled 180 upstaged and 180 pure DCIS eligible patients. Taking January 1, 2018 as the base point, we consecutively enrolled 120 pure DCIS before and 60 pure DCIS after. The same enrollment method was used for the upstaged patients. Data were collected between March 2016 and July 2018. The patients' ultrasound images were obtained from the hospital database. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the ethics committee of the principal investigator's hospital and is registered at ClinicalTrials.gov (050432-4-1911D). Because of the retrospective nature

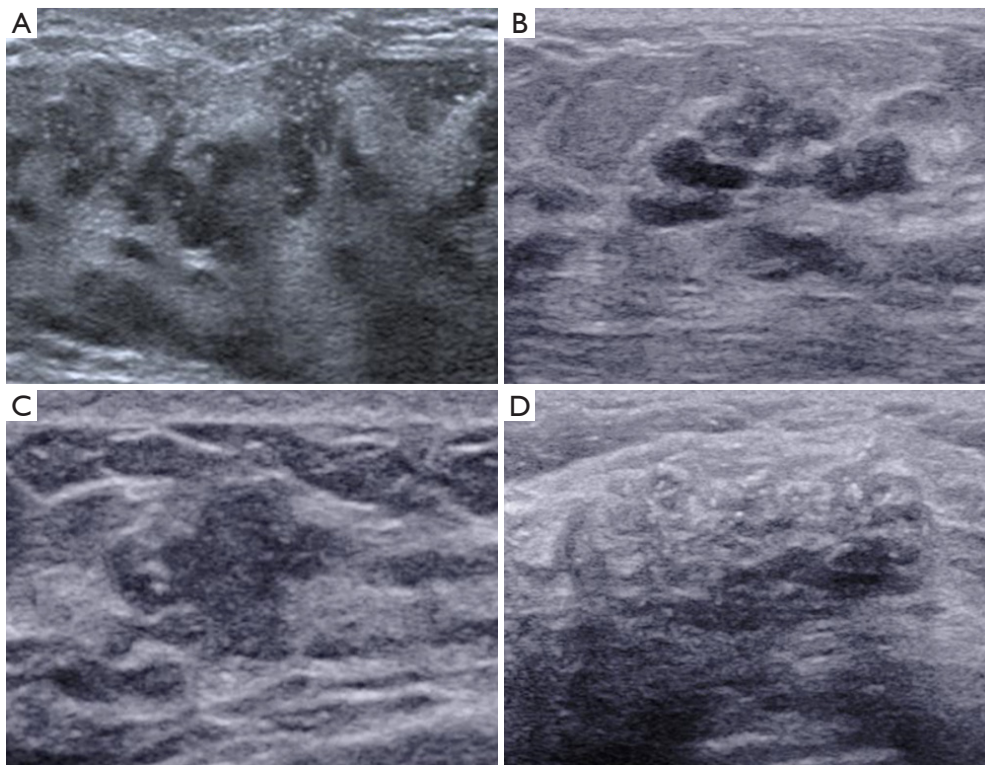


Figure 1 Complex and diverse ultrasound images of ductal carcinoma in situ. (A) Calcification is the main manifestation; (B) duct abnormalities are the main manifestation; (C) the mass is the main manifestation; (D) the structural disorder is the main manifestation.

of the research, the requirement for informed consent was waived.

The inclusion criteria were as follows: (I) the pathological diagnosis with CNB was pure DCIS; (II) the surgery was done at the Shanghai Cancer Hospital; (III) no adjuvant therapy, such as neoadjuvant chemotherapy, was performed before the operation; (IV) breast ultrasound examination was performed within a month before CNB, and the images were saved in the database. Patients were excluded from the study if the suspected anatomical sites based on the images did not have cancer on pathological analysis and if no obvious mass or non-mass lesion was detected by radiologists on ultrasound images.

We enrolled the patients consecutively. The first two-thirds of the patients in the two groups made up the training set (n=240), while the others made up the validation set (n=120).

Image acquisition and processing

The images in our research were from the Shanghai Cancer Hospital. DCIS is complex and has diverse ultrasound

images. Some lesions are diffusely distributed along the ducts (12,13). Therefore, it is difficult for radiologists to select a suitable region of interest (ROI) (see *Figure 1*). Therefore, the whole ultrasound image was taken as the ROI. Due to the variation in the image dimensions which were determined by the different ultrasound machines, all the images were resized into 200×200 pixels before putting them into the model. The label of each image was determined by the corresponding histopathological results; pure DCIS was 0, while upstaged DCIS was 1.

In our study, the models referred to the classical convolution neural network (CNN), including ResNet and VggNet (14-16). The classical model has been proven to be feasible in image feature recognition after many experiments. Therefore, our model retained their structure, and some adjustments were made to fit our data. In detail, as lesions occupy most of the images, and the training images were on a relatively small scale, we used fewer layers and changed the size of the frontier convolution kernel from 3×3 to 5×5. The output ranged from 0 to 1, indicating the probability of being upgraded.

Training set expansion was performed by mirror inversion

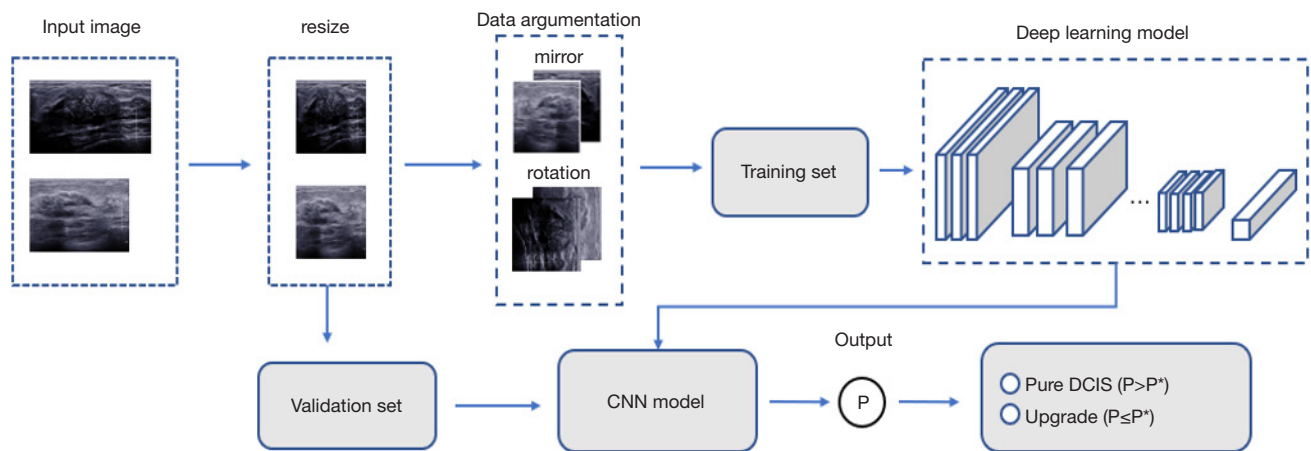


Figure 2 Schematic presentation of our proposal for classifying pure ductal carcinoma in situ (DCIS) and upgraded DCIS.

Table 1 Baseline characteristics of the patients

	Pure DCIS (n=180)	Upstaged DCIS (n=180)	%
Age			
≤50	66	93	44.16
>50	114	87	56.83
Tumor size on ultrasonography			
≤20	108	39	40.83
>20	72	141	59.16
Family history			
No	136	152	80.00
Yes	44	28	20.00
Menopause			
No	94	114	57.77
Yes	86	66	42.22

DCIS, ductal carcinoma in situ.

and rotation in multiples of 90°. The data expansion retained the image features and displayed features from different angles, which helped improve the robustness of the model and avoided overfitting. The training set, after expansion, was used to train the model, while the validation set was used to validate the performance of the model (Figure 2).

Due to the lack of test sets in this study, cross-validation was used to verify the stability and generalization of the model. Cross-validation can inhibit the sensitivity of the model to the data, so as to validate whether the model is stable.

We used a 3-fold cross-validation on the original data.

We randomly divided the two types of data into three parts, and took one part to constitute a validation set, while the others constituted the training set. Therefore, we had three combinations, and all of them were used to train and verify the performance of the model. Finally, we assessed the robustness of the model through the output.

We obtained the area under the receiver operating characteristic curve (AUROC), specificity, sensitivity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) of the model output for analysis.

Statistical analysis

We collected clinical data including age, family history, menopause, and tumor size on ultrasonography. A chi-square test was conducted to compare the clinical characteristics of the verification and test groups. A two-sided P value was used. P<0.05 was considered statistically significant. All statistical analyses were performed using SPSS version 25.0.

Results

The average age of the patients with pure DCIS was 54.9 years, and that of the upstaged patients was 49.9 years. The clinical data of the entire cohort is shown in Table 1. The training set included 240 patients, and the verification set included 120 patients. The ratio of upgraded DCIS to pure DCIS was 1:1 in both sets. Age, maximum tumor size on ultrasonography, family history, and menopause were not significantly different between the training set and the test set (P value >0.05) (Table 2).

Approximately 40% and 78% of “non-upgraded” and “upgraded” patients had a tumor size above 20 mm on the ultrasound image. Patients without a family history of upgraded DCI accounted for about 24%, and upgraded patients accounted for about 16%. Menopause patients accounted for about 48%, and upgraded patients accounted for approximately 37%. The age, maximum tumor size on ultrasonography, family history, and proportion of patients in menopause were comparable between the two groups (P value >0.05 for all).

Figure 3 and Table 3 show the results of the validation set

Table 2 Comparison of clinical features between the patients in the training set and validation set

	Training set (n=240)	Validation set (n=120)	Univariate P value
Age			0.178
≤ 50 years	105	54	
> 50 years	135	66	
Tumor size on ultrasonography			0.495
≤ 20	95	52	
> 20	145	68	
Family history			0.855
No	190	98	
Yes	50	22	
Menopause			0.097
No	146	62	
Yes	94	58	

and training set in the two types of models. Among ResNet models, the AUCROC of the validation set in the Resnet-b0 model was 0.804, with a sensitivity, specificity, accuracy, PPV, and NPV of 0.767, 0.716, 0.742, 0.730, and 0.754, respectively. The AUCROC of the validation set in the Resnet-b1 model was 0.821, with a sensitivity, specificity, accuracy, PPV, and NPV of 0.802, 0.733, 0.742, 0.746, and 0.738, respectively; the AUCROC of the validation set in the Resnet-b2 model was 0.737, with a sensitivity, specificity, accuracy, PPV, and NPV of 0.667, 0.683, 0.675, 0.678, and 0.672, respectively. In the Vgg-change model, the AUCROC of the validation set was 0.724, and the sensitivity, specificity, accuracy, PPV, and NPV were 0.717, 0.650, 0.683, 0.696, and 0.672, respectively. Considering the result of the training set (see Figure 3B), the performance between the training set and validation set was significantly different, which means that the Vgg-change model was overfitting and was not feasible for the data.

In the robustness verification experiments, a 3-fold cross-validation was used for all the models. The AUROCs of the 3-fold data sets in the ResNet-b0 model were 0.766, 0.817, and 0.738; they were 0.767, 0.808, and 0.760 in the ResNet-b1 model; and they were 0.759, 0.790, and 0.736 in the ResNet-b2 model (Figure 4). The performance of the ResNet-b1 model was the most feasible and stable.

Discussion

In this study, we established a deep learning model that uses two-dimensional ultrasound images to predict whether pure DCIS diagnosed by CNB will be upstaged postoperatively. The AUCROC of ResNet-b1 was 0.802, which is relatively stable. The accuracy and sensitivity of the validation set

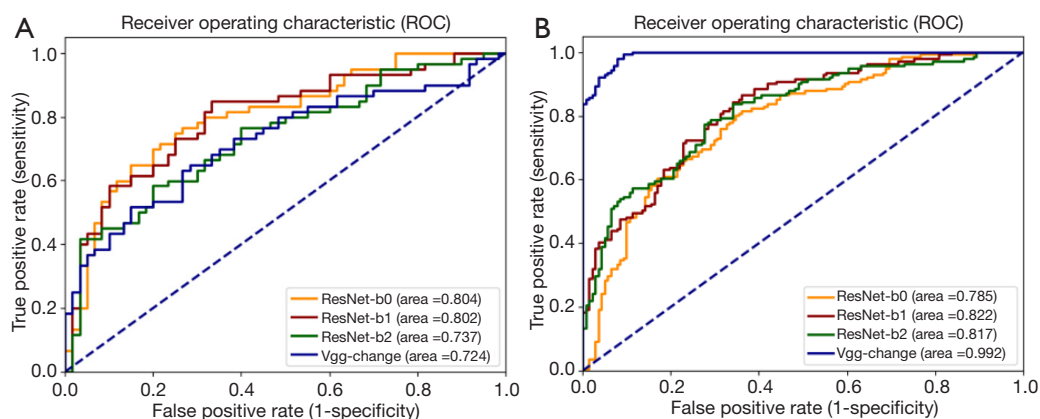


Figure 3 Receiver operating curve (ROC) for the four models. (A) ROC of the validation set; (B) ROC of the training set.

Table 3 Diagnostic performance of the deep learning algorithms for the test dataset

Algorithm	Validation		Sensitivity	Specificity	PPV	NPV
	AUROC	Accuracy				
ResNet-b0	0.804	0.742	0.767	0.716	0.730	0.753
ResNet-b1	0.802	0.742	0.733	0.750	0.745	0.738
ResNet-b2	0.737	0.675	0.667	0.683	0.678	0.672
Vgg-change	0.724	0.683	0.717	0.650	0.696	0.671

AUROC, area under the receiver operating characteristic curve; PPV, positive predictive value; NPV, negative predictive value.

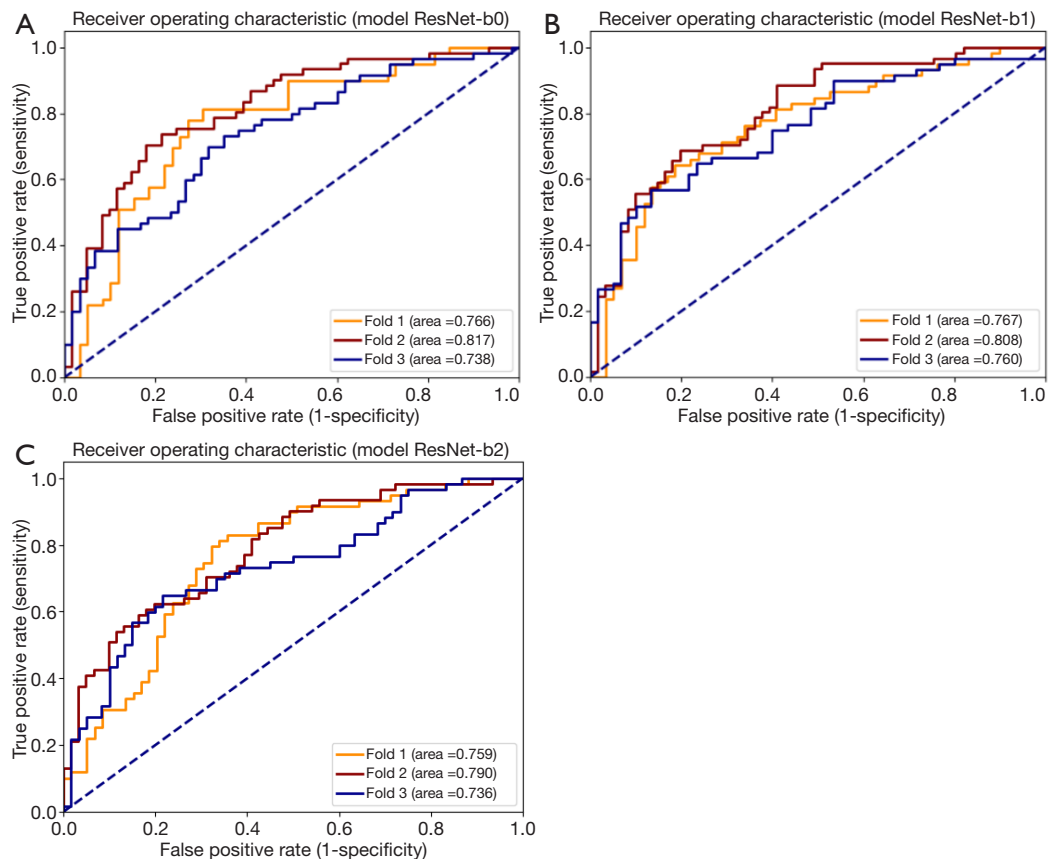


Figure 4 Comparison of robustness between the 3 deep learning models. (A) 3-fold cross-validation performance in the ResNet-b0 model; (B) 3-fold cross-validation performance in the ResNet-b1 model; (C) 3-fold cross-validation performance in the ResNet-b2 model.

were 74.2% and 73.3%, respectively. Our model can help surgeons decide whether SLNB should be performed.

Previous research has focused on exploring relevant clinical predictors (3,9,10,17,18). Multiple studies have reported clinical predictors for upstaging after CNB, such as age, size of the mass, and higher nuclear grade, and the relevant prediction models were established based on these factors. The AUCROC of previous models ranged from 0.58

to 0.70. Compared with these, the results of our model were better. Moreover, some relevant clinical factors are difficult to obtain in clinical practice, and evaluations of some of the factors are subjective. For example, in the James's model (19), the percentage of calcification remaining after CNB is relatively difficult to obtain, especially for non-calcified DCIS. In addition, whether the BI-RADS rating reaches 5 would be significantly affected by the radiologist's

experience. Although some clinical prediction models have a great AUC, they are difficult to apply in clinical practice. In the clinical prediction model developed by Coufal (20), although the AUCROCs reached 0.85, cases diagnosed as DCIS with microinvasion by CNB were included in the study. According to Champion *et al.* (21), although DCIS with microinvasion is a relatively special type between pure DCIS and invasive cancer, the current treatment and prognosis are closer to those of early invasive breast cancer. This model cannot accurately reflect the postoperative upstaging of pure DCIS diagnosed by CNB before surgery.

With the development of AI, the model can effectively integrate tumor image information and clinical information and transform it into an accurate clinical decision system. This is an important development direction for clinical adjuvant diagnosis and treatment in the future.

Compared with the traditional clinical models, AI is advantageous in that it can identify characteristic textures and details that radiologists cannot recognize, and it can quantitatively describe the image features, making its evaluation more objective.

To our knowledge, previous studies using AI to predict pure DCIS upgrades have been based on mammography or MRI images, and our study is the first to build a deep learning prediction model based on two-dimensional ultrasound images. In a study by Shi *et al.* (22), the researchers sketched suspicious lesions in mammography images and used the traditional machine learning method to let the computer learn the characteristics of the sketched suspicious lesions; its AUCROC was 0.70. Moreover, the ROI that was manually sketched by the radiologist was affected by their experience and subjective judgment. It is difficult to completely capture all image features of suspicious lesions; it is time- and labor-intensive. In the study by Mutasa *et al.* (23), although the method of deep learning was adopted to build a prediction model based on mammography images, its AUCROC was 0.71. In a study by Zhu *et al.* (24), they used MRI images as datasets of deep learning, but the AUCROC was 0.68 as well. Our AUCROC reached 0.802, which was relatively better compared with previous research on AI. In comparison, our adopted deep learning used the whole breast image as the ROI. As a result, the rich internal information based on large data from the entire image can provide better predictive models. Our method can also save time and effort. Compared with mammography, ultrasound has a more obvious advantage in evaluating the structural characteristics of impure calcifications (such as lumps and structural

distortions). In deep learning based on mammography (23), the specificity reached 92%, which is higher than ours. This may be because the sensitivity of ultrasound to focal calcifications is lower than that of mammography. However, it is noteworthy that the sensitivity of finding malignant calcifications on ultrasound is higher than that of finding benign calcifications (12).

According to relevant literature, only about 12–32% of pure DCIS diagnosed by CNB before surgery is upstaged to microinvasion or even to invasive cancer in postoperative pathology (3-6). This would result in an imbalance in the ratio of data between upgraded and pure DCIS, which might make the model's ability to diagnose upgraded DCIS weaker (11). In this study, two equal datasets were selected to reduce the bias of the model diagnosis and improve the robustness of the model.

This study has some limitations. First, this was a retrospective study. Data was acquired by different doctors using different ultrasound machines; therefore, the homogeneity of the data may be poor. Second, our study is a single-center study, which lacks an external verification set. To solve these existing limitations, we plan to conduct prospective studies in future to maintain uniformity of the images and to carry out multi-center cooperation to add external verification sets.

Conclusions

The AI model based on ultrasound images has a good and stable performance in predicting whether pure DCIS will be upgraded after verification in the verification group, and can provide guidance to clinicians when determining the surgical approach for DCIS.

Acknowledgments

Funding: This work was supported by the Ministry of Science and Technology of China under Grant No. 2017YFA0205200; the National Natural Science Foundation of China under Grants No. 61671449, 81227901, 81527805 and 81830058; the Science and Technology Commission of Shanghai Municipality under Grant No. 18411967400. Shanghai Municipal Commission of Health and Family Planning under grant No. 20174Y0011.

Footnote

Reporting Checklist: The authors have completed the

STARD 2015 reporting checklist. Available at <http://dx.doi.org/10.21037/atm-20-3981>

Data Sharing Statement: Available at <http://dx.doi.org/10.21037/atm-20-3981>

Peer Review File: Available at <http://dx.doi.org/10.21037/atm-20-3981>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-3981>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the ethics committee of the principal investigator's hospital and is registered at ClinicalTrials.gov (050432-4-1911D). Because of the retrospective nature of the research, the requirement for informed consent was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- DeSantis CE, Ma J, Goding Sauer A, et al. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J Clin* 2017;67:439-48.
- Lyman GH, Somerfield MR, Giuliano AE. Sentinel Lymph Node Biopsy for Patients With Early-Stage Breast Cancer: 2016 American Society of Clinical Oncology Clinical Practice Guideline Update Summary. *J Oncol Pract* 2017;13:196-8.
- Han JS, Molberg KH, Sarode V. Predictors of invasion and axillary lymph node metastasis in patients with a core biopsy diagnosis of ductal carcinoma in situ: an analysis of 255 cases. *Breast J* 2011;17:223-9.
- Sheaffer WW, Gray RJ, Wasif N, et al. Predictive factors of upstaging DCIS to invasive carcinoma in BCT vs mastectomy. *Am J Surg* 2019;217:1025-9.
- Mannu GS, Groen EJ, Wang Z, et al. Reliability of preoperative breast biopsies showing ductal carcinoma in situ and implications for non-operative treatment: a cohort study. *Breast Cancer Res Treat* 2019;178:409-18.
- Park AY, Gweon HM, Son EJ, et al. Ductal carcinoma in situ diagnosed at US-guided 14-gauge core-needle biopsy for breast mass: preoperative predictors of invasive breast cancer. *Eur J Radiol* 2014;83:654-9.
- Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer* 2015;51:2296-303.
- Hwang ES, Hyslop T, Lynch T, et al. The COMET (Comparison of Operative versus Monitoring and Endocrine Therapy) trial: a phase III randomised controlled clinical trial for low-risk ductal carcinoma in situ (DCIS). *BMJ Open* 2019;9:e026797.
- Doobar SC, de Monyé C, Stoop H, et al. Ductal carcinoma in situ diagnosed by breast needle biopsy: Predictors of invasion in the excision specimen. *Breast* 2016;27:15-21.
- Wiratkapun C, Patanajareet P, Wibulpholprasert B, et al. Factors associated with upstaging of ductal carcinoma in situ diagnosed by core needle biopsy using imaging guidance. *Jpn J Radiol* 2011;29:547-53.
- Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw* 2018;106:249-59.
- Wang LC, Sullivan M, Du H, et al. US appearance of ductal carcinoma in situ. *Radiographics* 2013;33:213-28.
- Ban K, Tsunoda H, Watanabe T, et al. Characteristics of ultrasonographic images of ductal carcinoma in situ with abnormalities of the ducts. *J Med Ultrason* (2001) 2020;47:107-15.
- Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *NeurIPS* 2012:1097-105.
- He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. *CVPR* 2016: 770-8.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*, 2014.
- Kondo T, Hayashi N, Ohde S, et al. A model to predict upstaging to invasive carcinoma in patients preoperatively diagnosed with ductal carcinoma in situ of the breast. *J Surg Oncol* 2015;112:476-80.

18. Lee SK, Yang JH, Woo SY, et al. Nomogram for predicting invasion in patients with a preoperative diagnosis of ductal carcinoma in situ of the breast. *Br J Surg* 2013;100:1756-63.
19. Jakub JW, Murphy BL, Gonzalez AB, et al. A Validated Nomogram to Predict Upstaging of Ductal Carcinoma in Situ to Invasive Disease. *Ann Surg Oncol* 2017;24:2915-24.
20. Coufal O, Selingerová I, Vrtělová P, et al. A simple model to assess the probability of invasion in ductal carcinoma in situ of the breast diagnosed by needle biopsy. *Biomed Res Int* 2014;2014:480840.
21. Champion CD, Ren Y, Thomas SM, et al. DCIS with Microinvasion: Is It In Situ or Invasive Disease? *Ann Surg Oncol* 2019;26:3124-32.
22. Shi B, Grimm LJ, Mazurowski MA, et al. Can Occult Invasive Disease in Ductal Carcinoma In Situ Be Predicted Using Computer-extracted Mammographic Features? *Acad Radiol* 2017;24:1139-47.
23. Mutasa S, Chang P, Van Sant EP, et al. Potential Role of Convolutional Neural Network Based Algorithm in Patient Selection for DCIS Observation Trials Using a Mammogram Dataset. *Acad Radiol* 2020;27:774-9.
24. Zhu Z, Harowicz M, Zhang J, et al. Deep learning analysis of breast MRIs for prediction of occult invasive disease in ductal carcinoma in situ. *Comput Biol Med* 2019;115:103498.

Cite this article as: Qian L, Lv Z, Zhang K, Wang K, Zhu Q, Zhou S, Chang C, Tian J. Application of deep learning to predict underestimation in ductal carcinoma in situ of the breast with ultrasound. *Ann Transl Med* 2021;9(4):295. doi: 10.21037/atm-20-3981