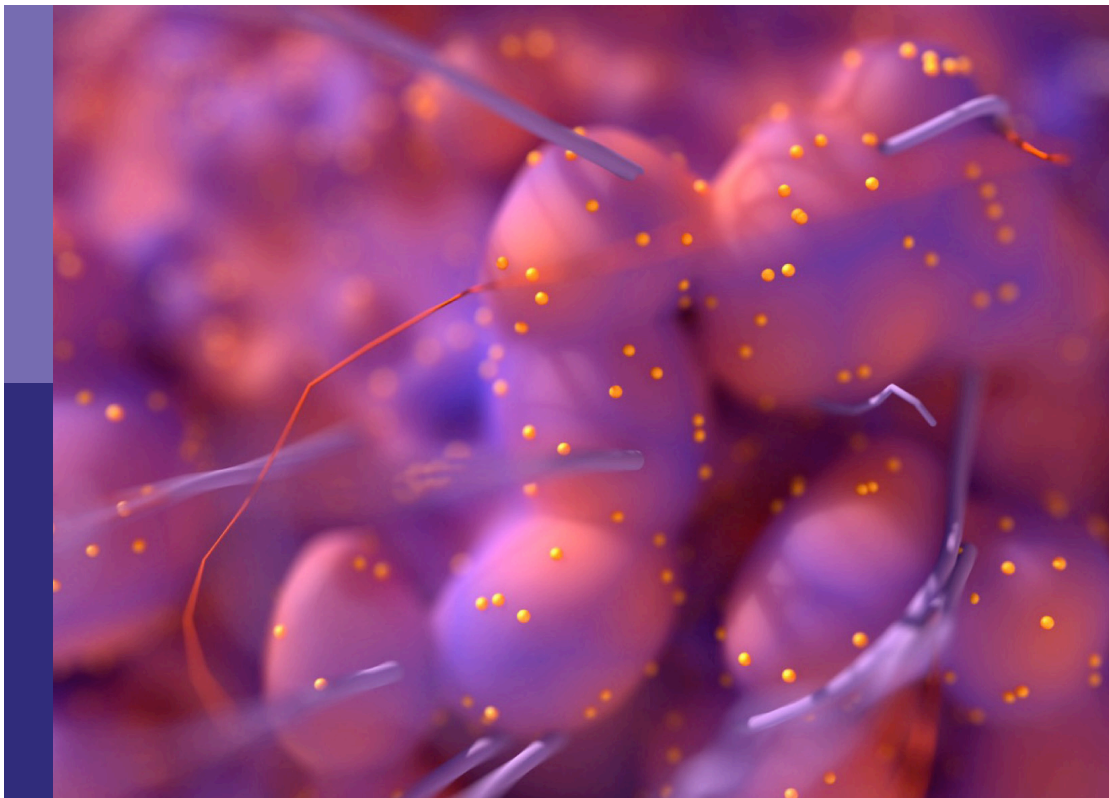
frontiers | Research Topics

Breast cancer imaging: clinical translation of novel methods

Edited by
Jiabao He and Luca Nicosia

Coordinated by
Sai Man Cheung and Simone Palma

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-8325-6204-8
DOI 10.3389/978-2-8325-6204-8

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Breast cancer imaging: clinical translation of novel methods

Topic editors
Jiabao He — Newcastle University, United Kingdom
Luca Nicosia — University of Insubria, Italy

Topic coordinators
Sai Man Cheung — Newcastle University, United Kingdom
Simone Palma — Agostino Gemelli University Polyclinic (IRCCS), Italy

Citation
He, J., Nicosia, L., Cheung, S. M., Palma, S., eds. (2025). *Breast cancer imaging: clinical translation of novel methods*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6204-8

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 20 March 2025
DOI 10.3389/fonc.2025.1581169

OPEN ACCESS

EDITED AND REVIEWED BY
Zaver Bhujwalla,
Johns Hopkins Medicine, United States

*CORRESPONDENCE
Sai Man Cheung
g.cheung@newcastle.ac.uk
Jiabao He
jiabao.he@newcastle.ac.uk

RECEIVED 21 February 2025
ACCEPTED 26 February 2025
PUBLISHED 20 March 2025

CITATION
Cheung SM, Palma S, Nicosia L and He J (2025) Editorial: Breast cancer imaging: clinical translation of novel methods. *Front. Oncol.* 15:1581169. doi: 10.3389/fonc.2025.1581169

Editorial: Breast cancer imaging: clinical translation of novel methods

Sai Man Cheung[1*], Simone Palma[2], Luca Nicosia[3] and Jiabao He[1*]

[1]Newcastle Magnetic Resonance Centre, Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom, [2]Department of Diagnostic Imaging, Oncological Radiotherapy and Hematology, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy, [3]Breast Imaging Division, Radiology Department, IEO European Institute of Oncology IRCCS, Milan, Italy

KEYWORDS

breast cancer imaging, clinical translation, MRI, ultrasound, early diagnosis, treatment monitoring

Editorial on the Research Topic:
Breast cancer imaging: clinical translation of novel methods

Breast cancer has become the most common cancer in women, and the incidence rate has increased by 6% in the last decade (1) with a projected increase of 2% between 2024 and 2035 (2). In the EU, women over 50 years of age receive regular radiological screening (3), while younger women at high risk of developing breast cancer receive annual surveillance (4). However, current radiological approaches are suboptimal and suffer from high false positive and negative rates (5), leading to overtreatment and late detection (6, 7). There is an urgent unmet clinical need for novel radiological methods to facilitate accurate early detection and treatment monitoring of breast cancer. Current radiological methods for breast cancer diagnosis and treatment monitoring are primarily mammography, ultrasound and MRI (8). Mammography is primarily sensitive to the presence of microcalcifications in the tumor, ultrasound is sensitive to solid masses in the tumor versus fluid-filled lesions, and MRI is sensitive to the presence of abnormal vasculature in the tumor (8). However, there have been major advances in medical imaging in recent years, ranging from novel ultrasound devices and algorithms to functional and metabolic MRI. These innovations not only have the potential to improve the accuracy of diagnosis but may also provide critical information for treatment planning that was previously unavailable from radiological examination, leading to a change in healthcare pathway. We, therefore, would like to highlight recent developments in breast imaging methods (6 articles on ultrasound techniques and 3 on MRI techniques) with this Research Topic to facilitate clinical translation.

Conventional B-mode ultrasound reconstructs breast anatomy from the reflection of high-frequency acoustic waves at the interface of tissue boundaries, offering a tool with the advantages of low cost, safety, speed, wide accessibility and high sensitivity in dense breast, and the disadvantages of limited image contrast and operator dependence (9). With a contrast agent to highlight blood flow, Li et al. explored the relationship of perfusion characteristics with molecular subtypes, and identified heterogeneous enhancement,

perfusion defects and peripheral radial vessels for grade III tumors, perfusion defects and clear edges after enhancement for human epidermal growth factor receptor 2 (HER-2) and triple-negative breast cancer (TNBC), and peak enhancement and wash-in perfusion for Luminal A and Luminal B differentiation. Using shear wave elastography to reveal mechanical properties and super microvascular imaging to outline microcirculation, Wang et al. investigated the inclusion of quantitative tumor properties in the breast-imaging reporting and data system (BI-RADS), and achieved higher sensitivity (+1.5%), specificity (+16.0%) and accuracy (+13.2%) than the conventional classification. Using strain elastography and an automated breast volume scanner to form a comprehensive picture, Shiyan et al. studied the risk of malignancy in hypoechoic lesions using a radiomics approach and constructed a nomogram using multivariate logistic regression with a larger area under the curve (AUC) in receiver operating characteristic (ROC) than the BI-RADS and clinical risk factors model alone. Leveraging artificial intelligence to accelerate workflow and reduce operator dependence (10), Qiu et al. trained a breast lesion classification algorithm using dynamic ultrasound videos from two hospitals and demonstrated a higher consistency closer to the experienced clinicians (Kappa: 0.82) than the junior clinicians (Kappa: 0.60) for diagnostic efficiency. Combining automated breast volume scanners and artificial intelligence to extract clinically relevant features, Li et al. estimated the probability of malignancy for ambiguous BI-RADS 4 lesions using radiomics features and showed an AUC of 0.949, a sensitivity of 82.14% and specificity of 95.56%. However, with RECIST criteria for treatment monitoring abandoned in many centers, Zhang et al. attempted to use an artificial intelligence algorithm trained on patients undergoing neoadjuvant chemotherapy (NACT) for pathological complete response identification but failed to show any significant improvement in AUC over manual and conventional approaches.

Conventional dynamic contrast-enhanced (DCE) MRI highlights the vascular abnormalities associated with angiogenesis in breast tumors using a paramagnetic contrast agent, offering a tool with the advantages of high resolution, high sensitivity, and good contrast in dense breast, and the disadvantages of limited image contrast, high cost, and potential adverse reactions to contrast agent (11). Using diffusion MRI for tissue microstructure profiling and an extension of intravoxel incoherent motion (IVIM) for concurrent microcirculation estimation without a contrast agent, Cheung et al. investigated tumor cellular microstructure and perfusion using a Bayesian algorithm for noise reduction as early response markers for NACT and found a decrease in perfusion fraction in good responders against an increase in poor responders after 1 cycle of NACT. However, with mathematical modeling of DCE MRI to derive quantitative perfusion characteristics, Almutlaq et al. explored the discrepancies in tumor perfusion quantified by IVIM and DCE MRI and illustrated the discordance between the two imaging techniques on perfusion but at the same time the concordance between interstitial and extracellular volume fractions against water diffusion. Using 2D, 3D and hotspot region-of-interest analysis approaches in conjunction with quantitative metrics of apparent diffusion coefficient to reduce operator dependency, Biswas et al. derived the optimal cutoffs with 2 and 4 diffusion weightings using two larger clinical trials [ECOG-ACRIN A6702 (12) and EUSOBI (13)] and recommended the hotspot approach with 2 diffusion weightings for differentiating between malignant and benign lesions.

We thank all the contributors for their excellent research work that advances medical imaging for the diagnosis and prognosis of breast cancer. Together, we can help humankind identify breast cancer early and treat it more gently.

Author contributions

SMC: Writing – original draft, Writing – review & editing. SP: Writing – review & editing. LN: Writing – review & editing. JH: Writing – original draft, Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

References

1. Barclay NL, Burn E, Delmestri A, Duarte-Salles T, Golozar A, Man WY, et al. Trends in incidence, prevalence, and survival of breast cancer in the United Kingdom from 2000 to 2021. *Sci Rep*. (2024) 14:19069. doi: 10.1038/s41598-024-69006-1

2. Smittenaar CR, Petersen KA, Stewart K, Moitt N. Cancer incidence and mortality projections in the UK until 2035. *Br J Cancer*. (2016) 115:1147–55. doi: 10.1038/bjc.2016.304

3. Loibl S, André F, Bachelot T, Barrios CH, Bergh J, Burstein HJ, et al. Electronic address: clinicalguidelines@esmo.org. Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol*. (2024) 35:159–82. doi: 10.1016/j.annonc.2023.11.016

4. Sessa C, Balmaña J, Bober SL, Cardoso MJ, Colombo N, Curigliano G, et al. Electronic address: clinicalguidelines@esmo.org. Risk reduction and screening of cancer in hereditary breast-ovarian cancer syndromes: ESMO Clinical Practice Guideline. *Ann Oncol*. (2023) 34:33–47. doi: 10.1016/j.annonc.2022.10.004

5. Nelson HD, O'Meara ES, Kerlikowske K, Balch S, Miglioretti D. Factors associated with rates of false-positive and false-negative results from digital mammography screening: an analysis of registry data. *Ann Intern Med*. (2016) 16164:226–35. doi: 10.7326/M15-0971

6. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JM, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer*. (2015) 51:2296–303. doi: 10.1016/j.ejca.2015.07.017

7. Nguyen DL, Shelley Hwang E, Ryser MD, Grimm LJ. Imaging changes and outcomes of patients undergoing active monitoring for ductal carcinoma in situ: seven-year follow-up study. *Acad Radiol*. (2024) 31:2654–62. doi: 10.1016/j.acra.2023.12.021

8. Barba D, León-Sosa A, Lugo P, Suquillo D, Torres F, Surre F, et al. Breast cancer, screening and diagnostic tools: All you need to know. *Crit Rev Oncol Hematol*. (2021) 157:103174. doi: 10.1016/j.critrevonc.2020.103174

9. Shen S, Zhou Y, Xu Y, Zhang B, Duan X, Huang R, et al. A multi-centre randomised trial comparing ultrasound vs mammography for screening breast cancer in high-risk Chinese women. *Br J Cancer*. (2015) 112:998–1004. doi: 10.1038/bjc.2015.33

10. Moy L. Change is good: the evolution and future of breast imaging. *Radiology*. (2023) 306:e230018. doi: 10.1148/radiol.230018

11. Kuhl CK. MRI of breast tumors. *Eur Radiol*. (2000) 10:46–58. doi: 10.1007/s003300050006

12. Rahbar H, Zhang Z, Chenevert TL, Romanoff J, Kitsch AE, Hanna LG, et al. Utility of diffusion-weighted imaging to decrease unnecessary biopsies prompted by breast MRI: A trial of the ECOG-ACRIN cancer research group (A6702). *Clin Cancer Res*. (2019) 25:1756–65. doi: 10.1158/1078-0432.CCR-18-2967

13. Baltzer P, Mann RM, Iima M, Sigmund EE, Clauser P, Gilbert FJ, et al. Diffusion-weighted imaging of the breast-a consensus and mission statement from the EUSOBI International Breast Diffusion-Weighted Imaging working group. *Eur Radiol*. (2020) 30:1436–50. doi: 10.1007/s00330-019-06510-3

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 05 September 2023
DOI 10.3389/fonc.2023.1192630

OPEN ACCESS

EDITED BY
Luca Nicosia,
European Institute of Oncology (IEO), Italy

REVIEWED BY
Bilgin Kadri Aribas,
Bülent Ecevit University, Türkiye
Ottavia Battaglia,
Università degli Studi di Milano, Italy

*CORRESPONDENCE
Liangyu Wang
290499799@qq.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 23 March 2023
ACCEPTED 22 August 2023
PUBLISHED 05 September 2023

CITATION
Wang X, He Y and Wang L (2023) Diagnostic value of shear wave elastography combined with super microvascular imaging for BI-RADS 3-5 nodules.
Front. Oncol. 13:1192630.
doi: 10.3389/fonc.2023.1192630

Diagnostic value of shear wave elastography combined with super microvascular imaging for BI-RADS 3-5 nodules

Xueqing Wang†, Yi He† and Liangyu Wang*

Department of Ultrasound, Shantou Central Hospital, Shantou, Guangdong, China

Background: To investigate the diagnostic value of shear wave elastography (SWE) and super microvascular imaging (SMI) integrated with the traditional ultrasound breast imaging reporting and data system (BI-RADS) classification in differentiating between benign and malignant breast nodules.

Methods: For analysis, 88 patients with 110 breast nodules assessed as BI-RADS 3-5 by conventional ultrasound were selected. SWE and SMI evaluations were conducted separately, and all nodules were verified as benign or malignant ones by pathology. Receiver operating characteristic (ROC) curves were plotted after obtaining quantitative parameters of different shear waves of nodules, including maximum (Emax), mean (Emean), minimum (Emin) Young's modulus, modulus standard deviation (SD), and modulus ratio (Eratio). The best cut-off value, specificity, sensitivity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) for diagnosing malignant nodules employing Emax were obtained, and the diagnostic value of combining Emax and BI-RADS classification was compared. SMI graded nodule based on the Alder blood flow grading standard, whereas the BI-RADS classification was based on microvascular morphology. We assessed the diagnostic value of SMI for breast nodules and investigated the diagnostic efficacy of SWE combined with SMI in differentiating benign and malignant breast nodules with BI-RADS classification 3–5.

Results: The adjusted the BI-RADS classification using SMI and SWE technologies promoted the sensitivity, specificity, and accuracy of discriminating benign and malignant breast nodules ($P < 0.05$). The combination of traditional ultrasound BI-RADS classification with SWE and SMI technologies offered high sensitivity, specificity, accuracy, PPV, and NPV for identifying benign and malignant breast lesions. Moreover, combining SWE and SMI technologies with the adjusted BI-RADS classificationhad the best diagnostic efficacy for distinguishing benign and malignant breast nodules with BI-RADS 3–5.

Conclusion: The combination of SWE and SMI with the adjusted BI-RADS classification is a promising diagnostic method for differentiating benign and malignant breast nodules.

KEYWORDS

breast cancer, breast nodules, shear wave elastography, super microvascular imaging, microvascular morphology classification

1 Introduction

In clinical practice, imaging approaches including ultrasound, mammography, CT, and MRI are often employed for breast cancer screening. Mammography has been the preferred choice for early screening of breast cancer in the past (1–3). However, the sensitivity of mammography for detecting dense breast nodules is low (4, 5). Traditional ultrasound evaluation has become the preferred screening approach for the early detection of breast nodules because of its simplicity, safety, noninvasiveness, radiation-free, and real-time dynamic imaging benefits. However, breast nodules have various appearances on grayscale ultrasound images, and benign and malignant features overlap; therefore, the application of new ultrasound technologies is urgently required to enhance diagnostic efficacy (6–9).

SWE is a non-invasive, real-time, dynamic ultrasound imaging technology that evaluates tissue hardness by measuring the velocity of shear wave propagation in the target tissue. The principle is that the propagation speed of the mechanical wave is proportional to the hardness of the propagating medium, and the ultrasonic probe forms a continuous ultrasonic shear wave source in the target region, and then the propagation speed of the shear wave is accurately measured to calculate the hardness of the propagating medium. The relationship between shear wave velocity and microstructure hardness is positive, and the higher the velocity, the greater the microstructure hardness.Since malignant nodules are harder than benign ones, studies have demonstrated that the average shear wave velocity of malignant lesions is substantially higher than that of benign lesions (10–12). Breast tumors comprise various tissue components, primarily consisting of tumor cells and surrounding stromal components. In the growth process of malignant breast tumors, tumor cells constantly proliferate, infiltrate, necrose, and repair, resulting in collagen synthesis and fibrous tissue proliferation and reactive proliferation of surrounding connective tissue (13, 14). The higher the density of tumor cells, the more edema in the surrounding tissue and the higher the hardness of the tumor. This forms the pathological histological basis for employing elastography to differentiate between benign and malignant breast nodules (15, 16).

SMI technology is a real-time non-invasive microvascular imaging technology, which can detect low-velocity microvessels with high resolution, high frame rate and minimum motion artifacts. It adopts a unique adaptive algorithm to eliminate clutter and motion artifacts generated by source tissue motion through a multi-dimensional wall filter, thus minimizing the loss of low-speed blood flow information. Compared to traditional blood flow imaging (such as color Doppler and power Doppler), which can only show vessels with higher flow velocity and tube diameter > 0.2mm, SMI can visualize low-velocity tiny vessels with tube diameter > 0.1mm without injecting contrast agent. It represents a technological innovation in vascular imaging and is mainly used for the assessment of tumor vessels (17). Investigations have reported that it can substantially enhance the diagnostic effectiveness of benign and malignant breast nodules (18). The growth, invasion, and metastasis of breast cancer are closely associated with the formation of new microvessels. Therefore, the emergence and use of ultramicro vascular imaging technology offer a good complement to traditional ultrasound blood flow detection methods. It can not only measure low-speed microvessels with diameters greater than 0.1 mm but can also effectively separate low-flow signals from tissue motion artifacts, even retaining the finest low-flow components.

SWE and super microvascular imaging (SMI) technologies were combined in this research to reclassify breast imaging reporting and data system (BI-RADS) categories based on the hardness information and blood flow signal characteristics of breast nodules and to further examine the diagnostic value of the two new technologies for distinguishing between benign and malignant breast nodules classified as BI-RADS 3-5.

2 Materials and methods

2.1 Study population

Patients who visited the Central Hospital of Shantou fbecause of breast nodules were chosen as the study population. Inclusion criteria were as follows: (a) BI-RADS classification, SWE, and SMI diagnostic examinations were conducted with no contraindications; (b) complete imaging data were available. Exclusion criteria were as follows: (1) patients with a history of breast surgery, breast cancer recurrence, or concomitant malignant tumors; (2) lesions larger than 4 cm in maximum diameter or with internal liquefaction; (3) patients receiving neoadjuvant chemotherapy; (4) patients with a history of breast implantation; (5) lactating or pregnant women; (6) patients with mental or cognitive disorders. All lesions had undergone biopsy or surgical pathological diagnosis. 88 patients with 110 lesions were included, with ages ranging from 15 to 87 years and a mean age of (46.39 ± 14.81) years. The maximum diameter of breast lesions ranged from 4 to 40 mm, with a mean of (19.34 ± 7.96) mm.

2.2 Instruments and methods

2.2.1 Instruments

A Japanese TOSHIBA Aplioi900 ultrasound diagnostic system equipped with SWE and SMI imaging technology with a high-frequency linear array probe of 10–14 MHZ was used. The "BREAST" mode of the instrument system settings was chosen. An ultrasound physician with over 10 years of experience and proficiency in SWE and SMI technologies guided the operator, and three operations were repeated to obtain an average value and reduce human error.

2.2.2 Methods

First, a two-dimensional (2D) gray-scale breast ultrasound evaluation was conducted. The ultrasound probe was gently placed on the breast and radially scanned from the nipple. This process was repeated twice. The depth, gain, and focus were

adjusted according to the lesion condition after detecting the breast lesion to obtain the best image quality. Sonographic features of the breast lesion were recorded in detail, and any abnormal lymph nodes in the axilla were also screened.

2.2.3 Classification criteria and imaging observation indicators

Breast BI-RADS classification: The comprehensive analysis was conducted in accordance with the classification criteria of the 2013 version of the Breast Imaging Reporting and Data System (BI-RADS). Class 3: High probability of benign or low grade malignancy, the probability of malignancy is 0%-2%, short-term follow-up of 3 to 6 months; Class 4: Suspected malignancy, the possibility of malignancy is 3%-94%, biopsy is recommended; Class 5: Highly suggestive of malignancy, malignant probability ≥ 95%, recommended biopsy and active treatment. BI-RADS class 4 lesions were further classified into 4a, 4b, and 4c subcategories (18–20), as follows: 4a - low suspicion for malignancy, with a 3-10% likelihood of malignancy; 4b - intermediate suspicion for malignancy, with a 10-50% likelihood of malignancy; 4c - high suspicion for malignancy, with a 51-94% likelihood of malignancy. Malignant signs, such as microcalcifications, irregular shape, spiculated margins, were categorized, round shape, microlobulated/indistinct/angular margins, duct extension, complex echogenicity and posterior acoustic shadowing, non-parallel growth, were categorized (19, 20).

SMI: The number, course, and distribution of microvessels inside and around the lesion are observed using CDFI and mSMI technology after determining the location of the breast lesion using routine ultrasound. The size of the sampling box was modified, including regulating the blood flow velocity measurement range within 1 cm around the nodule and its surrounding breast tissue as much as possible, which is approximately 1.0–2.0 cm/s. To classify the microvascular morphology of breast nodules in SMI mode, Adler's blood flow grading standard (21) was employed (22): (1) avascular type: no visible blood flow signal was detected within the nodule; (2) linear type: a single linear or slightly curved blood flow signal was detected within the nodule without crossing; (3) branching type: blood flow signals with uniform vessel diameter and branching were detected within the nodule, similar to branching; (4) root type: the blood vessel course within the nodule is irregular and disordered, and less than two large twisted blood vessels can be detected around it; (5) crab claw type: two or more radiating, thick and twisted blood vessels, or tiny, thorn-like blood vessels can be detected around the nodule. The microvascular morphology distribution was classified based on the above types after obtaining the mSMI blood flow image of the benign and malignant breast nodules (23). Among them, nodules with microvascular morphology types of avascular, linear, and branching were judged as benign nodules, and those with residual root and crab claw types were judged as malignant nodules.

SWE examination: SWE mode was initiated by the same physician using the same ultrasound diagnostic equipment after verifying the location of the lesion using traditional ultrasound (24–26). The measuring range was set to 0–180 kPa. The ROI was drawn to include the entire nodule using a grayscale ultrasound display of the lesion boundary. The SWE elasticity average value (Emean) of the lesion was measured. Then, the ROI range was set to 2 mm × 2 mm and placed in the elasticity mode map to obtain different SWE parameters including Emax, Emin, Eratio, SD, and Emean. Each data was measured three times and the average value was taken.

2.3 Statistics

For analysis, SPSS19.0 software was employed. Count data were presented as mean ± standard deviation ($x \pm s$), whereas t-tests and chi-square tests were employed for continuous and categorical data, respectively. The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and diagnostic accuracy for the diagnosis of breast nodules employing the BI-RADS classification, SWE, and SMI technology alone, and the two technologies combined with the BI-RADS classification diagnostic criteria, were separately computed. To compare the diagnostic indicators among different methods, chi-square tests were employed. Furthermore, receiver operating characteristic (ROC) curves were separately constructed for SWE, SMI, BI-RADS classification, and their combinations, and the area under the curve (AUC) was computed. The optimal cutoff value was determined as the elastic value with the maximum Youden index. A statistically significant difference was considered when $P < 0.05$.

3 Results

3.1 Pathological results

In this research, pathological findings were obtained from 110 solid nodules in 88 patients, which were either obtained through biopsy or surgical excision (Table 1). Of these, 60 nodules (54.54%) were malignant and 50 nodules (45.45%) were benign (Table 1).

3.2 Conventional ultrasound BI-RADS classification results

The traditional ultrasound evaluation reported that among the 110 breast nodules, 26 (23.63%) nodules were categorized as BI-RADS 3 and were all benign; 14 (12.72%) nodules were categorized as BI-RADS 4a, with 2 (1.81%) nodules being malignant and 12 (10.90%) nodules being benign; 27 (24.54%) nodules were categorized as BI-RADS 4b, with 17 (15.45%) nodules being malignant and 10 (9.09%) nodules being benign; 32 (29.09%) nodules were categorized as BI-RADS 4c, with 30 (27.27%) nodules being malignant and 2 (1.81%)nodules being benign; 11 (10%)nodules were categorized as BI-RADS 5 and were all malignant. BI-RADS 3-4a lesions were considered benign, while BI-RADS 4b-5 lesions were considered malignant. There were two cases of infiltrating ductal carcinoma among the misdiagnosed

TABLE 1 Pathological results of 110 breast nodules.

Pathological Results	Number (n)	Percentage (%)
Benign	50	
Fibroadenoma	37	33.63
Breast adenosis	5	4.54
Benign phyllodes tumor	3	2.72
Intraductal papilloma	3	2.72
Breast inflammation	2	1.81
Malignant	60	
Invasive ductal carcinoma	48	43.63
Invasive lobular carcinoma	3	2.72
Ductal carcinoma in situ	4	3.63
Mucinous carcinoma	3	2.72
Medullary carcinoma	1	0.90
Invasive small cell carcinoma	1	0.90
Total	110	

malignant nodules. There were six cases of fibroadenoma, four cases of mammary gland hyperplasia, one case of chronic mastitis, and one case of lymphocytic mastitis among the misdiagnosed benign lesions.

3.3 Examination results of BI-RADS classification combined with shear wave elastography technology in routine ultrasound

3.3.1 Shear wave elastography evaluation of breast nodules

The Young's modulus values of breast nodules obtained through SWE measurement, including Emax, Emean, Emin, and SD, demonstrated substantial differences between the benign and malignant groups ($P < 0.01$) (Table 2).

3.3.2 Diagnostic performance of combined use of BI-RADS and shear wave elastography in breast nodule evaluation

In this research, BI-RADS classification and SWE were combined, with Emax ≥77.25 kPa and BI-RADS classification greater than 4a employed as the malignant standard. ROC curves were constructed for Emax (SWE), BI-RADS classification (US), and the combined diagnostic approach (US+SWE) (Table 3, Figure 1). The sensitivity of US+SWE was not substnatially different from that of US or SWE alone, while its NPV was slightly higher than that of the two individual classifications. The specificity, accuracy, and PPV of US and SWE substantially increased with the combined US+SWE approach. A statistically significant difference was observed in the AUC between the US+SWE approach and the US alone ($Z = 3.404$, $P = 0.0007$).

3.4 Results of the examination using traditional ultrasound BI-RADS classification combined with super microvascular imaging technology

3.4.1 Alder grading of breast nodule microvessels in SMI mode

In SMI mode, the blood flow classification of benign nodules based on the Alder grading system (Table 4) demonstrated statistically significant differences ($\chi^2 = 44.153$, $P<0.01$) compared with the pathological results. The microvascular morphology of benign nodules tended to be avascular [10 (19.4%)], linear [19 (38.00%)], or dendritic [19 (38.00%)], whereas malignant nodules tended to be root-like [20 (33.33%)] or crab-like [25 (41.67%)]. In SMI mode, there was a significant difference in microvascular morphology between benign and malignant breast nodules ($\chi2 = 56.181$, $P<0.01$) (Table 4).

3.4.2 Optimization and adjustment of BI-RADS classification using SMI technology

In SMI mode, the BI-RADS classification remained unchanged or was downgraded when the microvascular morphology of breast nodules was classified as benign (avascular, linear, or dendritic). The BI-RADS classification remained unchanged or was upgraded when the microvascular morphology was classified as malignant (root-like or crab-like). The nodule classification remained unchanged when the low-level grayscale ultrasound showed a low-level microvascular classification and the high-level grayscale ultrasound showed a high-level microvascular classification. The classification was upgraded by one level when the low-level grayscale ultrasound showed a high-level microvascular classification. However, the classification could be downgraded by at most one level when the high-level grayscale ultrasound showed a low-level microvascular classification. Table 5 shows a comparison of the adjusted BI-RADS classification with the pathological results.

TABLE 2 Comparison of SWE elasticity modulus parameters between benign and malignant breast lesions.

Pathology Result	Number	Emax/kPa	Emean/kPa	Emin/kPa	SD/kPa
Benign	50	42.95 ± 37.34	26.60 ± 21.95	15.69 ± 11.85	4.80 ± 6.04
Malignant	60	114.28 ± 23.83	66.40 ± 19.38	28.31 ± 11.86	13.39 ± 9.86
t value		12.133	10.099	5.555	5.369
P value		<0.001	<0.001	<0.001	<0.001

TABLE 3 Diagnostic performance of various examination methods for differentiating benign and malignant breast nodules.

Examination Method	Examination Result	Pathological Result (Number of Nodules)		Sensitivity (%)	Specificity (%)	Accuracy (%)	Positive Predictive Value (%)	Negative Predictive Value (%)	AUC
		Benign	Malignant						
US	Benign	38	2	96.67	76.00	82.27	82.86	95.00	0.938
	Malignant	12	58						
US+SWE	Benign	46	2	96.67	84.00	90.91	87.88	95.45	0.975
	Malignant	4	58						
US+SMI	Benign	46	3	95.00	92.00	93.63	93.44	93.88	0.966
	Malignant	4	57						
US+SWE+SMI	Benign	46	1	98.33	92.00	95.45	97.87	96.72	0.980
	Malignant	4	59						

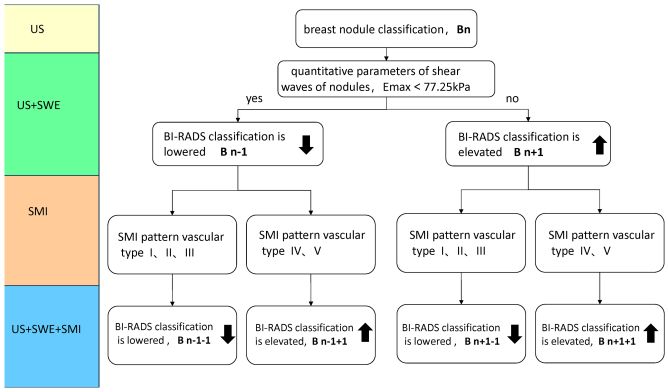


FIGURE 1
Classification Rule (Bn) for incorporating traditional ultrasound BI-RADS with SWE or SMI in Figure 1. When BI-RADS 3, Bn classification, n=1; BI-RADS-US 4a, Bn classification, n=2; BI-RADS 4b, Bn classification, n=3; BI-RADS 4c, Bn classification, n=4; BI-RADS 5, Bn classification, n=5. For instance, a nodule was initially classified as BI-RADS 3. Its Bn classification is 1, when Emax > 77.25kPa, it is classified as B2 (n+1) by adding SWE to traditional ultrasound (US + SWE); when the SMI mode is type III branch, it is classified as B1 (n+1-1), by adding SMI on US + SWE (US + SWE + SMI), Bn classification ranges from 1 to 5 after elevated and lowered. (SMI pattern vascular type: type I avascular, type II linear, type III branching, type IV root-like, type V crab-li.

TABLE 4 Microvascular morphology classification of breast nodules under SMI mode [n (%)].

Pathological Type	Number	avascular	linear	dendritic	root-like	crab-like
Benign	50	10 (20)	19 (38)	19 (38)	1 (2)	1 (2)
Malignant	60	3 (5)	3 (5)	9 (15)	20 (33.33)	25 (41.67)
Total	110	13	22	28	21	26

3.4.3 Diagnostic performance of the adjusted BI-RADS classification for benign and malignant breast nodules

Compared with the traditional ultrasound BI-RADS classification (US) after adjustment, the adjusted BI-RADS classification employing SMI microvascular morphology classification (US+SMI) demonstrated a slightly lower sensitivity and NPV (P<0.05) (Table 3, Figure 1). However, compared with the prior adjustment, the specificity, accuracy, and PPV were significantly improved ($\chi 2$ = 4.763, P<0.05). The AUC for US+SMI was 0.966, and there was a statistically significant difference in the AUC between US+SMI and US (Z = 2.826, P = 0.0047). Combining SMI classification can enhance the diagnostic accuracy of benign and malignant breast nodules classified as BI-RADS 3-5.

TABLE 5 Comparison analysis of BI-RADS classification combined with SMI adjustment and pathological results for breast nodules (n).

BI-RADS Category	Nodules(n)		Pathological Result Benign		Pathological Result Malignant		Diagnostic Accuracy of Malignant Nodules (%)	
	Before Adjust	After Adjust	Before Adjust	After Adjust	Before Adjust	After Adjust	Before Adjust	After Adjust
3	26	37	26	37	0	0	0.00	0.00
4a	14	12	12	9	2	3	14.29	25.00
4b	27	19	10	2	17	17	62.96	89.47
4c	32	27	2	2	30	25	93.75	92.59
5	11	15	0	0	11	15	100.00	100.00
Total	110	110	50	50	60	60		

3.5 The examination results of the traditional ultrasound BI-RADS classification combined with the SWE and SMI are concluded based on the following criteria

The BI-RADS classification is elevated or unchanged when Emax is greater than or equal to 77.25 kPa and/or the mSMI pattern is malignant vascular type (IV-type root-like or V-type crab-like). The BI-RADS classification is lowered or unchanged when Emax is less than 77.25 kPa and/or the SMI pattern is benign vascular type (I type avascular, II type linear, or III type branching) (Figure 1).

Results showed that the AUC values of US+SWE+SMI, US+SWE, and US+SMI were 0.980, 0.975, and 0.966, respectively. Among them, the US+SWE+SMI had the highest AUC and diagnostic values (Table 3, Figure 2) compared to that of US-BI-RADS classification ($P < 0.01$). No statistical difference was observed in the AUC value between US+SWE+SMI and US+SWE or US+SMI. For distinguishing between benign and malignant breast nodules, the sensitivity, specificity, accuracy, PPV, and NPV of US+SWE+SMI were all superior to those of US, US+SWE, or US+SMI.

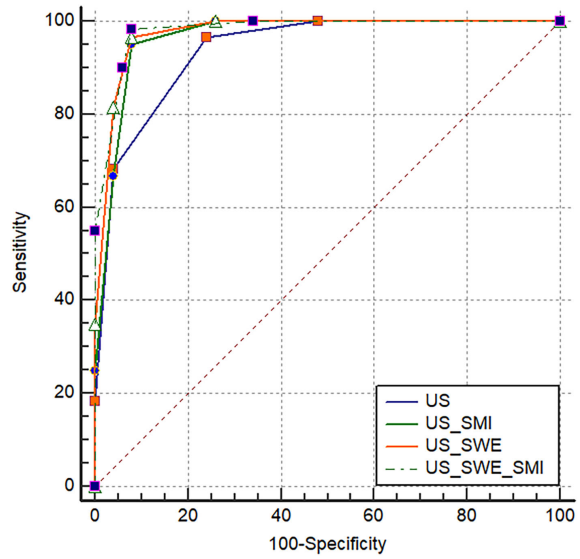


FIGURE 2
Receiver operating characteristic curves for US and three sets of combined tests.

4 Discussion

To re-adjust BI-RADS classification based on the nodule hardness and blood flow characteristics of breast nodules, this research combined SWE and SMI technologies and further examined the diagnostic value of the two new technologies for discriminating benign and malignant BI-RADS 3-5 breast nodules (exmple Figures 3, 4). This research discovered that the sensitivity, specificity, and accuracy of the BI-RADS classification adjusted by SMI and SWE technologies for the diagnosis of breast nodules were higher than those before adjustment ($P<0.05$). Combining SWE and SMI with traditional ultrasound BI-RADS classification can improve the diagnostic efficiency of BI-RADS 3-5 breast nodules, and their combination has the highest diagnostic efficiency, offering a more reliable diagnostic basis for clinical practice.

4.1 Classification ultrasound BI-RADS results

BI-RADS 3-4a nodules were classified as benign lesions and 4b-5 nodules were classified as malignant lesions based on the conventional ultrasound BI-RADS classification criteria in this experiment. The sensitivity of the traditional ultrasound BI-RADS classification for the diagnosis of benign and malignant breast nodules was high, but the specificity was low, indicating a high misdiagnosis rate of traditional ultrasound BI-RADS classification. Due to that, a large proportion of benign breast nodules will be misdiagnosed as malignant one. Some of the misdiagnosed benign lesions were fibroadenomas with active proliferation of surrounding ductal epithelium, and some fibroadenomas were accompanied by surrounding glandular disease, resulting in the misdiagnosis as malignant. Misdiagnosed malignant nodules had parallel growth, uniform internal echogenicity, slightly blurred margins, apparent capsules, and slight attenuation of posterior echoes; therefore, they

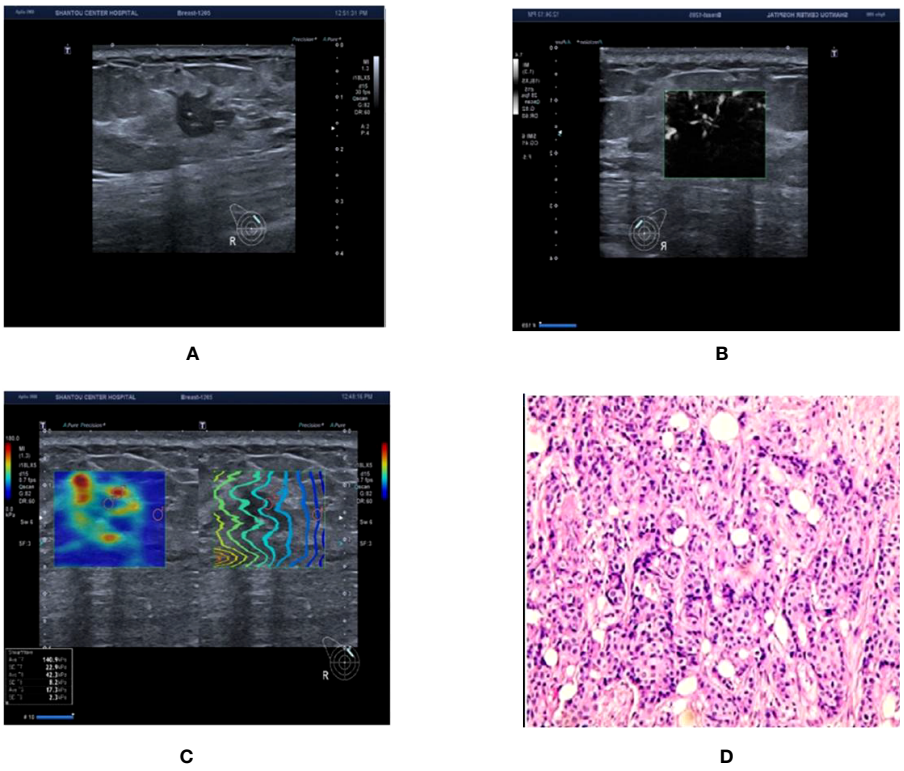


FIGURE 3
Ultrasound and pathological graphs: A 50-year-old female patient with a right breast nodule received ultrasound and pathological biopsy examinations. **(A)** An irregularly shaped, crab-like hypoechoic nodule is detected in the upper outer quadrant of the right breast, with unclear boundaries, classified as BI-RADS 4c. **(B)** mSMI image demonstrates that the microvascular morphology is radicle-shaped, with an increased BI-RADS level of 1 classified as BI-RADS 5. **(C)** Shear wave elastography image: Emax is 140.9 kPa, which is higher than 77.25 kPa, resulting in an increase or no change in the BI-RADS classification. **(D)** The nodule is an invasive ductal carcinoma based on pathological results.

were categorized as BI-RADS 4a. Particularly in BI-RADS 4, the malignancy rate of breast nodules previously examined using ultrasound as BI-RADS 3-5 varies significantly, and there were several similar morphological signs between benign and malignant nodules. Although diagnosed strictly according to the BI-RADS classification criteria, it is still challenging to accurately determine their classification sometimes. Furthermore, many patients with benign breast nodules have undergone excessive biological examination and surgery due to the misdiagnosis, which is unnecessary (27). Thus, it is urgent to develop diagnostic methods with high accuracy to reduce unnecessary clinical intervention.

4.2 Diagnostic performance of BI-RADS classification adjusted with SWE

This research combined SWE technology and BI-RADS classification criteria to exclude false-positive diagnoses of eight lesions in the BI-RADS classification standard. In cases where the lesion was small and the degree of fibrosis in the breast cancer cell matrix was low, the collagen fiber content decreased, resulting in a corresponding decrease in the hardness (28, 29). It has been reported that the higher breast thickness and lesion depth can result in lower elasticity values compared with actual values during SWE examination (19). Our results revleaed that the combination of US and SWE had an improved specificity and comparable sensitivity compared to US or SWQ alone. Hence, the addition of nodule elasticity contributes to the high diagnostic accuracy.

4.3 Diagnostic performance of BI-RADS classification adjusted with SMI

The blood flow levels of benign tumors were mostly at levels 0–1, while those of malignant tumors were mostly at levels 2–3. Malignant tumors have an abundant blood supply and a higher blood flow rate than benign tumors. Park et al. discovered that SMI was superior to color Doppler when evaluating the morphology, quantity and distribution of tumor microvessels (30). However, some fibroadenomas with large volume or quick growth rates will have similar vascular blood flow signals and penetrating vessel patterns compared to malignant tumors when using semiquantitative grading and penetrating vessels as diagnostic criteria (31, 32). Therefore, employing blood flow distribution pattern analysis as a diagnostic criterion is superior to semiquantitative grading in diagnosing benign and malignant breast lesions (33).

In this research, Adler grade 0-1 nodules with avascular, linear, dendritic subtype of microvascular morphology were classified as benign ones, while Adler grade 2-3 nodules with root and crab leg subtype of microvascular morphology were classified as malignant

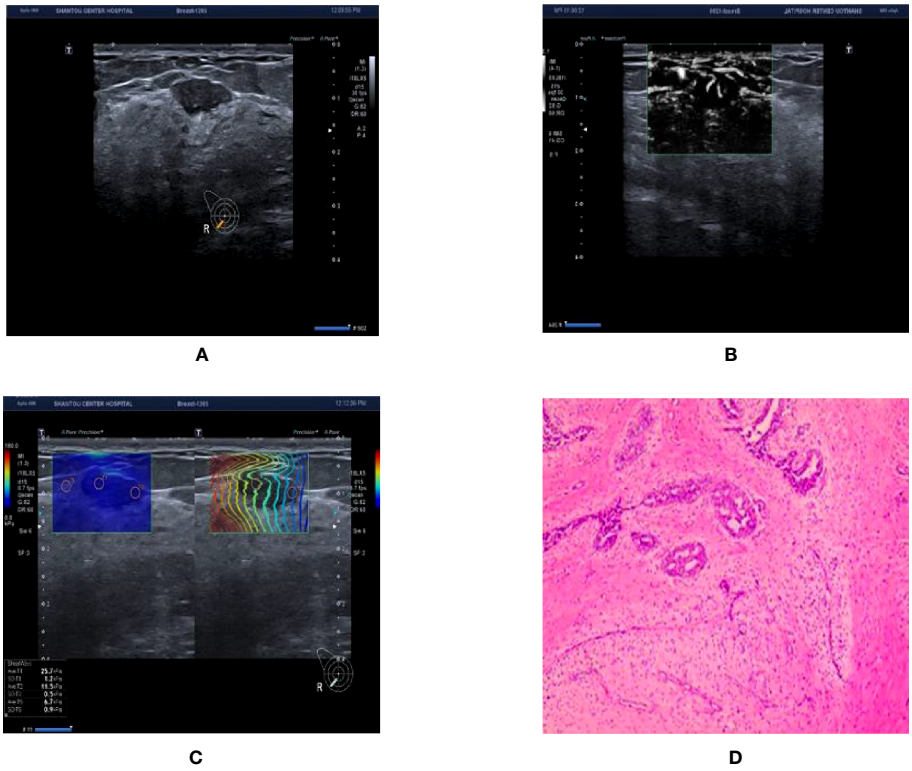


FIGURE 4
Ultrasound and pathological graphs. A 28-year-old female patient with a right breast nodule received ultrasound and pathological biopsy examinations. **(A)** A regular low-echo nodule was detected in the right outer and lower mammary quadrant, with parallel growth, lobed edges, even internal echo, and clear boundaries, BI-RADS 4a. **(B)** mSMI shows that microvessel morphology is dendritic subtype; BI-RADS grade was downgraded 1 level. **(C)** Shear wave elastic imaging: Emax is 25.7kPa < 77.25 kPa; BI-RADS classification was lowered or remained unaltered. **(D)** The nodule is a breast fibroadenoma based on pathological results.

ones (34, 35). A significant difference of microvascular morphology classification and the Adler grade was observed between the benign and malignant breast nodules ($P < 0.01$). Furthermore, when we adjusted the BI-RADS classification in combination with SMI technology, four benign nodules were classified as malignant, two of which were classified as 4b type, both of which were sclerosing adenosis with poor blood supply. The other two cases were classified as 4c nodules, which were lymphocytic mastitis and chronic granulomatous mastitis, respectively, and both of them exhibited obvious malignant signs on traditional ultrasound with the Adler grade of 3 and the microvascular morphology of root and crab leg subtypes under the mSMI mode. Three malignant nodules were categorized as benign ones (BI-RADS 4a), of which two were verified to be ductal carcinoma in situ, and the other was verified to be infiltrating lobular carcinoma. Since traditional ultrasound only showed slight malignant signs, so downgrading blindly is not advisable. A total of 18 BI-RADS classifications of breast nodules were upgraded, and no downgrade was adopted. The combination of SMI and BI-RADS classification criteria can enhance the diagnostic accuracy of breast nodules but still has limitation.

In this research, 4b nodules showed avascular characteristics and were downgraded to 4a. Postoperative pathology demonstrated that they were breast glandular diseases with fibroadenoma. BI-RADS 4a nodules exhibited linear characteristics and were downgraded to category 3. Postoperative pathology indicated that they were fibroadenoma with peripheral proliferation. All two cases of 2D ultrasound images demonstrated irregular shapes and unclear borders. Among them, breast glandular disease with fibroadenoma was more common in clinical cases, with nodules frequently having unclear borders and a hard texture, which can be categorized as BI-RADS category 4. Six 4a nodules downgraded to grade 3 were all fibroadenomas, which were routinely classified as 4a because of their irregular morphology, some with angular or lobulated shapes but with a benign vascular pattern. According to the pathological findings, most of these patients were middle-aged to elderly women, and the nodules had been present in their bodies for numerous years without substantial changes in size. Pathological findings demonstrated that chronic inflammation around the edge of the nodules could result in irregular nodular morphology, and the blood flow distribution inside the nodules was not abundant. Therefore, the combination of SMI and traditional BI-RADS classification can help differentiate these modules.

4.4 Diagnostic performance of BI-RADS classification optimized by combining SWE and SMI technologies

In this research, 15 nodules were SWE-positive but SMI-negative, whereas 2 nodules were SWE-negative but SMI-positive.

48 nodules were positive for both technologies. When the findings were inconsistent, the BI-RADS category remained unchanged. The addition of SWE and SMI to the BI-RADS classification led to 26 fibroadenomas remaining unaltered, 9 fibroadenomas categorized as BI-RADS 4a being downgraded to level 3, and 8 benign 4b nodules being downgraded, indicating that not all level 4 nodules require immediate intervention or biopsy since they were benign nodules. However, five nodules were still misdiagnosed even after combining the SWE and SMI approaches.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Ethics Committee of Shantou Central Hospital. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin because Approved by the Ethics Council.

Author contributions

LW and XW designed the project. XW and YH drafted the manuscript. XW performed the experiments. XW and YH performed the data analysis. YH and LW critically reviewed and revised the manuscript. LW approved the submitted version of the manuscript. XW and YH contributed equally to this work.

Funding

This work was supported from the Shantou medical health science and technology project.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Lee SH, Ryu HS, Jang MJ, Yi A, Ha SM, Kim SY, et al. Glandular tissue component and breast cancer risk in mammographically dense breasts at screening breast US. *Radiology* (2021) 301(1):57–65. doi: 10.1148/radiol.2021210367

2. Schünemann HJ, Lerda D, Quinn C, Follmann M, Alonso-Coello P, Rossi PG, et al. Breast cancer screening and diagnosis: A synopsis of the European breast guidelines. *Ann Internal Med* (2020) 172(1):46–56. doi: 10.7326/M19-2125

3. Yaffe MJ, Jong RA. Adjunctive ultrasonography in breast cancer screening. *Lancet* (2016) 387(10016):313–4. doi: 10.1016/S0140-6736(15)00787-4

4. Ciatto S, Visioli C, Paci E, Zappa M. Breast density as a determinant of interval cancer at mammographic screening. *Br J cancer.* (2004) 90(2):393–6. doi: 10.1038/sj.bjc.6601548

5. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *Jama-journal Am Med Assoc* (2008) 299 (18):2151–63. doi: 10.1001/jama.299.18.2151

6. Hall FM. Mammographic density of the breast. *New Engl J Med* (2003) 348 (2):174–5. doi: 10.1056/NEJM200301093480215

7. Houssami N, Ciatto S. Ultrasound and mammography for breast cancer screening. *Jama-journal Am Med Assoc* (2008) 300(13):1514–5. doi: 10.1001/jama.300.13.1514-a

8. Dogan BE, Turnbull LW. Imaging of triple-negative breast cancer. *Ann Oncol* (2012) 23 Suppl 6(null):vi23–9. doi: 10.1007/978-3-319-69980-6_4

9. Britton P, Sinnatamby R. Investigation of suspected breast cancer. *BMJ (Clinical Res ed).* (2007) 335(7615):347–8. doi: 10.1136/bmj.39234.386470.be

10. Zhou J, Zhan W, Chang C, Zhang J, Yang Z, Dong Y, et al. Role of acoustic shear wave velocity measurement in characterization of breast lesions. *J ultrasound Med* (2013) 32(2):285–94. doi: 10.7863/jum.2013.32.2.285

11. Spick C, Baltzer PA. Diagnostic utility of second-look US for breast lesions identified at MR imaging: systematic review and meta-analysis. *Radiology* (2014) 273 (2):401–9. doi: 10.1148/radiol.14140474

12. Krouskop TA, Wheeler TM, Kallel F, Garra BS, Hall T. Elastic moduli of breast and prostate tissues under compression. *Ultrasonic imaging.* (1998) 20(4):260–74. doi: 10.1177/016173469802000403

13. Haas JS, Kaplan CP. The divide between breast density notification laws and evidence-based guidelines for breast cancer screening: legislating practice. *JAMA Internal Med* (2015) 175(9):1439–40. doi: 10.1001/jamainternmed.2015.3040

14. Orel S. Who should have breast magnetic resonance imaging evaluation? *J Clin Oncol* (2008) 26(5):703–11. doi: 10.1200/JCO.2007.14.3594

15. Yang WT, Hennessy BT, Dryden MJ, Valero V, Hunt KK, Krishnamurthy S. Mammary angiosarcomas: imaging findings in 24 patients. *Radiology* (2007) 242 (3):725–34. doi: 10.1148/radiol.2423060163

16. Park SY, Park JY, Park JW, Kim WH, Park JY, Kim HJ. Unexpected hyperechoic lesions of the breast and their correlations with pathology: a pictorial essay. *Ultrasonography* (2022) 41(3):597–609. doi: 10.14366/usg.21243

17. Chen SC, Cheung YC, Su CH, Chen MF, Hwang TL, Hsueh S. Analysis of sonographic features for the differentiation of benign and Malignant breast tumors of different sizes. *Ultrasound obstetrics gynecology.* (2004) 23(2):188–93. doi: 10.1002/uog.930

18. Xiao XY, Chen X, Guan XF, Wu H, Qin W, Luo BM. Superb microvascular imaging in diagnosis of breast lesions: a comparative study with contrast-enhanced ultrasonographic microvascular imaging. *Br J Radiol* (2016) 89(1066):20160546. doi: 10.1259/bjr.20160546

19. Yoon JH, Jung HK, Lee JT, Ko KH. Shear-wave elastography in the diagnosis of solid breast masses: what leads to false-negative or false-positive results? *Eur Radiol* (2013) 23(9):2432–40. doi: 10.1007/s00330-013-2854-6

20. de Margerie-Mellon C, Debry JB, Dupont A, Cuvier C, Giacchetti S, Teixeira L, et al. Nonpalpable breast lesions: impact of a second-opinion review at a breast unit on BI-RADS classification. *Eur radiology.* (2021) 31(8):5913–23. doi: 10.1007/s00330-020-07664-1

21. Adler DD, Carson PL, Rubin JM, Quinn-Reid D. Doppler ultrasound color flow imaging in the study of breast cancer: preliminary findings. *Ultrasound Med Biol* (1990) 16(6):553–9. doi: 10.1016/0301-5629(90)90020-D

22. Liu B, Zheng Y, Huang G, Lin M, Shan Q, Lu Y, et al. Breast lesions: quantitative diagnosis using ultrasound shear wave elastography-A systematic review and meta–analysis. *Ultrasound Med Biol* (2016) 42(4):835–47. doi: 10.1016/j.ultrasmedbio.2015.10.024

23. Du J, Li FH, Fang H, Xia JG, Zhu CX. Microvascular architecture of breast lesions: evaluation with contrast-enhanced ultrasonographic micro flow imaging. *J ultrasound Med* (2008) 27(6):833–42. doi: 10.7863/jum.2008.27.6.833

24. Xiao X, Jiang Q, Wu H, Guan X, Qin W, Luo B. Diagnosis of sub-centimetre breast lesions: combining BI-RADS-US with strain elastography and contrast-enhanced ultrasound-a preliminary study in China. *Eur radiology*. (2017) 27 (6):2443–50. doi: 10.1007/s00330-016-4628-4

25. Mitka M. New ultrasound "elasticity" technique may reduce need for breast biopsies. *Jama-journal Am Med Assoc* (2007) 297(5):455. doi: 10.1001/jama.297.5.455

26. Mitka M. New screening methods offer hope for more accurate breast cancer detection. *Jama-journal Am Med Assoc* (2008) 299(4):397–8. doi: 10.1001/jama.2008.2

27. Lam WW, Chu WC, Tse GM, Ma TK. Sonographic appearance of mucinous carcinoma of the breast. *Am J roentgenology*. (2004) 182(4):1069–74. doi: 10.2214/ajr.182.4.1821069

28. Bai M, Zhang HP, Xing JF, Shi QS, Gu JY, Li F, et al. Acoustic radiation force impulse technology in the differential diagnosis of solid breast masses with different sizes: which features are most efficient? *BioMed Res Int* (2015) 2015(null):410560. doi: 10.1155/2015/410560

29. Evans A, Whelehan P, Thomson K, McLean D, Brauer K, Purdie C, et al. Quantitative shear wave ultrasound elastography: initial experience in solid breast masses. *Breast Cancer Res* (2010) 12(6):R104. doi: 10.1186/bcr2787

30. Park AY, Seo BK, Cha SH, Yeom SK, Lee SW, Chung HH. An innovative ultrasound technique for evaluation of tumor vascularity in breast cancers: superb micro-vascular imaging. *J Breast Cancer*. (2016) 19(2):210–3. doi: 10.4048/jbc.2016.19.2.210

31. Schroeder RJ, Bostanjoglo M, Rademaker J, Maeurer J, Felix R. Role of power Doppler techniques and ultrasound contrast enhancement in the differential diagnosis of focal breast lesions. *Eur radiology*. (2003) 13(1):68–79. doi: 10.1007/s00330-002-1413-3

32. Cosgrove DO, Kedar RP, Bamber JC, al-Murrani B, Davey JB, Fisher C, et al. Breast diseases: color Doppler US in differential diagnosis. *Radiology* (1993) 189(1):99–104. doi: 10.1148/radiology.189.1.8372225

33. Yongfeng Z, Ping Z, Wengang L, Yang S, Shuangming T. Application of a novel microvascular imaging technique in breast lesion evaluation. *Ultrasound Med Biol* (2016) 42(9):2097–105. doi: 10.1016/j.ultrasmedbio.2016.05.010

34. Wang Q, Li XL, He YP, Alizad A, Chen S, Zhao CK, et al. Three-dimensional shear wave elastography for differentiation of breast lesions: An initial study with quantitative analysis using three orthogonal planes. *Clin hemorheology microcirculation*. (2019) 71(3):311–24. doi: 10.3233/CH-180388

35. Lichtenbeld HC, Barendsz-Janson AF, van Essen H, Struijker Boudier H, Griffioen AW, Hillen HF. Angiogenic potential of Malignant and non-Malignant human breast tissues in an in *vivo* angiogenesis model. *Int J cancer*. (1998) 77 (3):455–9. doi: 10.1002/(SICI)1097-0215(19980729)77:3<455::AID-IJC23>3.0.CO;2-5

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 17 October 2023
DOI 10.3389/fonc.2023.1256146

OPEN ACCESS

EDITED BY
Luca Nicosia,
European Institute of Oncology (IEO), Italy

REVIEWED BY
Dania Cioni,
University of Pisa, Italy
Zhiyong Zhou,
Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Zhang Yan
546940642@qq.com

RECEIVED 10 July 2023
ACCEPTED 27 September 2023
PUBLISHED 17 October 2023

CITATION
Shiyan G, Liqing J, Yueqiong Y and Yan Z (2023) A clinical-radiomics nomogram based on multimodal ultrasound for predicting the malignancy risk in solid hypoechoic breast lesions.
Front. Oncol. 13:1256146.
doi: 10.3389/fonc.2023.1256146

A clinical-radiomics nomogram based on multimodal ultrasound for predicting the malignancy risk in solid hypoechoic breast lesions

Guo Shiyan, Jiang Liqing, Yan Yueqiong and Zhang Yan*

Department of Ultrasound, Third Xiangya Hospital, Central South University, Changsha, Hunan, China

Background: In routine clinical examinations, solid hypoechoic breast lesions are frequently encountered, but accurately distinguishing them poses a challenge. This study proposed a clinical-radiomics nomogram based on multimodal ultrasound that enhances the diagnostic accuracy for solid hypoechoic breast lesions.

Method: This retrospective study analyzed ultrasound strain elastography (SE) and automated breast volume scanner images (ABVS) of 423 solid hypoechoic breast lesions from 423 female patients in our hospital between August 2019 and May 2022. They were assigned to the training (n=296) and validation (n=127) groups in a 7:3 ratio by generating random numbers. Radiomics features were extracted and screened from ABVS and SE images, followed by the calculation of the radiomics score (Radscore) based on these features. Subsequently, a nomogram was constructed through multivariate logistic regression to assess the malignancy risk in breast lesions by combining Radscore with Breast Imaging Reporting and Data System (BI-RADS) scores and clinical risk factors associated with breast malignant lesions. The diagnostic performance, calibration performance, and clinical usefulness of the nomogram were assessed by the area under the curve (AUC) of the receiver operating characteristic curve, the calibration curve, and the decision analysis curve, respectively.

Results: The diagnostic performance of the nomogram is significantly superior to that of both the clinical diagnostic model (BI-RADS model) and the multimodal radiomics model (SE+ABVS radiomics model) in training (AUC: 0.972 vs 0.930 vs 0.941) and validation group (AUC:0.964 vs 0.916 vs 0.933). In addition, the nomogram also exhibited a favorable goodness-of-fit and could lead to greater net benefits for patients.

Conclusion: The nomogram enables a more effective assessment of the malignancy risk of solid hypoechoic breast lesions; therefore, it can serve as a new and efficient diagnostic tool for clinical diagnosis.

KEYWORDS

nomogram, breast, radiomics, automated breast volume scanner, strain elastography

1 Introduction

As the most prevalent cancer in the world, breast cancer poses a grave threat to people's health and survival (1). Given its high metastatic tendency and high mortality rate (2, 3), coupled with the significant differences in treatment modalities for benign and malignant breast tumors, early definitive diagnosis is a critical first step in the therapeutic management of breast lesions, which plays a crucial role in improving patient outcomes and survival (3–5).

With the recent advancements in ultrasound imaging technology, ultrasound plays an increasingly important role in the detection of breast lesions. Strain elastography (SE) allows for a quick and intuitive display of differences in elasticity coefficients within the lesion through color-coded imaging, therefore, it serves as a powerful diagnostic aid to offer valuable reference values for lesion diagnosis (6, 7). Automated breast volume scanner (ABVS) provides good reproducibility of diagnostic results due to its standardized operating procedures (8), it can acquire the whole breast volume information and perform multiplanar imaging on the acquired information. Studies have shown that ABVS exhibits comparable diagnostic accuracy to handheld ultrasound scanners for detecting breast lesions, while also providing additional information (9, 10). In routine ultrasound examinations, it is frequent to encounter patients with solid hypoechoic breast lesions, physicians can make an initial assessment of the malignancy risk of breast lesions based on their morphological appearance on ABVS image and elastic performance on SE images. The combination of ABVS and SE imaging techniques demonstrates significant diagnostic efficacy in evaluating breast lesions (11). However, the dependability of diagnostic outcomes generated by conventional imaging techniques is largely contingent on the proficiency of the examining physician and is markedly susceptible to interobserver variability (12).

Radiomics is in line with the current trend toward precision medicine, as it transforms ordinary visual images into high-throughput data through deep mining of medical images, allowing for the capturing of the internal heterogeneity of the entire tumor in a non-invasive manner (13–15). Therefore, it may provide novel biomarkers to facilitate diagnosis for better clinical decision-making. There are already several radiomics studies on ultrasound (US), mammography, and magnetic resonance (MR) in breast cancer diagnosis that have yielded promising results (16–29). However, there have been no studies on the combination of ABVS and UE radiomics features with clinical ultrasound factors for the diagnosis of breast cancer. Therefore, we conducted a radiomics analysis on SE and ABVS images, then combined these features with traditional imaging risk assessments and other clinical risk factors, resulting in a novel nomogram to help physicians accurately diagnose solid hypoechoic breast lesions.

2 Materials and methods

2.1 Patients

The retrospective study was approved by the institutional review board at our hospital. The inclusion criteria were: (1) Patients who underwent both ABVS and SE examinations at our hospital between August 2019 and May 2022 and subsequently underwent biopsy or surgical resection within two weeks with a pathologically confirmed diagnosis. (2) Patients with complete imaging data. (3) Patients' breast lesions were hypoechoic solid lesions. The exclusion criteria were: (1) Patients whose images were of poor quality; (2) Patients who underwent aspiration or clinical treatment before examining target lesions. Eventually, we included 423 solid hypoechoic breast lesions from 423 female patients. By generating random numbers, they were allocated into training and validation groups in the ratio of 7:3. The flow is shown in Figure 1.

2.2 Image acquisition and assessment

In this study, all images used were obtained using the ACUSON S2000 US machine and its accompanying ABVS system.

2.2.1 SE image

Patients were instructed to breathe normally while lying supine on the examination bed with their breasts fully exposed. A 9L4 probe in two-dimensional ultrasound mode was used to examine the breast in all planes. The imaging mode was then switched to the elastic mode when scanning the largest two-dimensional section of the lesions, and the patient was required to cooperate by holding her breath. The probe is placed perpendicularly over the breast without applying any pressure during the capture of SE images, with the lesion positioned at the center of an elasticity sampling window at least twice the size of the area of interest.

2.2.2 ABVS image

Instruct the patient to raise both arms over the head and remain in the supine position. A sufficient amount of coupling agent was

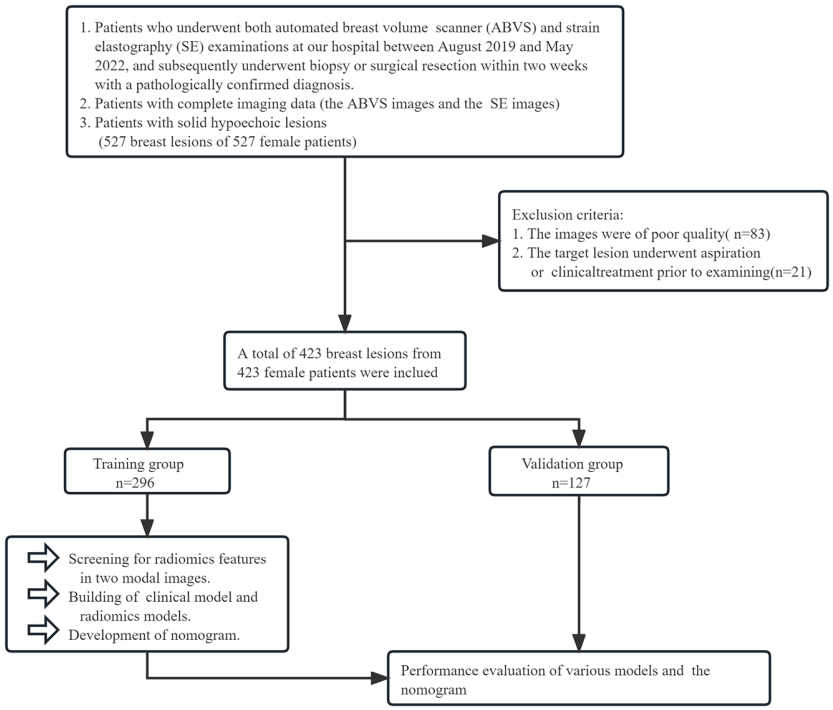


FIGURE 1
The grouping process of this study.

applied uniformly to the breast. Before scanning, parameters such as depth and overall gain were adjusted to achieve optimal image quality. During the scanning procedure, patients were instructed to breathe normally. Each breast was routinely scanned in two positions using a 14L5BV high-frequency linear array automatic scanning US probe. The nipples were marked after the scanning, and the acquired images were saved and transferred to a workstation for processing and analysis. Images with the maximum section image of the target lesion in coronal, transverse, and sagittal planes were selected for subsequent region of interest (ROI) segmentation and feature extraction.

Refer to the BI-RADS criteria which were defined by the American College of Radiology in 2013 (30), we evaluated the morphology, margin, border, orientation, posterior echogenicity of the lesion, microcalcification within the lesion, and conditions of retraction in its coronal plane on the saved images. This is followed by a combination with ultrasound elastic strain performance of the lesions (elasticity score using a 5-point scale (31)), enabling an accurate classification of the risk of malignancy of the lesions.

All of the above steps were performed by a physician who has over ten years of expertise in ultrasound breast disease diagnosis.

2.3 Extraction and selection of radiomics features

The ABVS and SE images were sequentially imported into 3D Slicer 5.2.1 for image processing, manual segmentation of the ROI and extraction of radiomics features. We delineated ROI for lesions in both SE and ABVS images. Notably, in the ABVS images, we delineated the ROI on the coronal, sagittal, and transverse planes of the lesions, respectively. Further details can be found in Figure 2. This procedure was performed with the participation of two physicians. Physician A, with five years of experience in ultrasound-based breast disease diagnosis, performed outlining for all the lesions. Physician B, with eight years of experience in ultrasound-based breast disease diagnosis, conducted lesion outlining on the training group to validate ROI outlining reproducibility. Then, we utilized the Pyradiomics package within 3D Slicer to extract radiomics features from the SE and ABVS images, respectively. The features extracted encompassed first-order statistics features, texture features (including the gray level co-occurrence matrix (glcm), gray level dependence matrix (gldm), gray level run length matrix (glrlm), gray level size zone matrix (glszm), and the neighbouring gray tone difference matrix (ngtdm)), as well as post-wavelet transformed features. Subsequently, we subjected the extracted features to screening. By utilizing the intra-class correlation coefficient (ICC) analysis, we can identify features that exhibit high levels of reproducibility (ICC > 0.75) (32), subsequently, radiomics features extracted from the region of interest segmented by physician A were utilized for further analysis. All feature values were normalized using Zscore. The radiomics features of both modalities were subjected to dimensionality reduction through the Mann-Whitney U test and least absolute shrinkage and selection operator (LASSO) regression and to identify features with strong qualitative diagnostic ability for solid hypoechoic breast lesions.

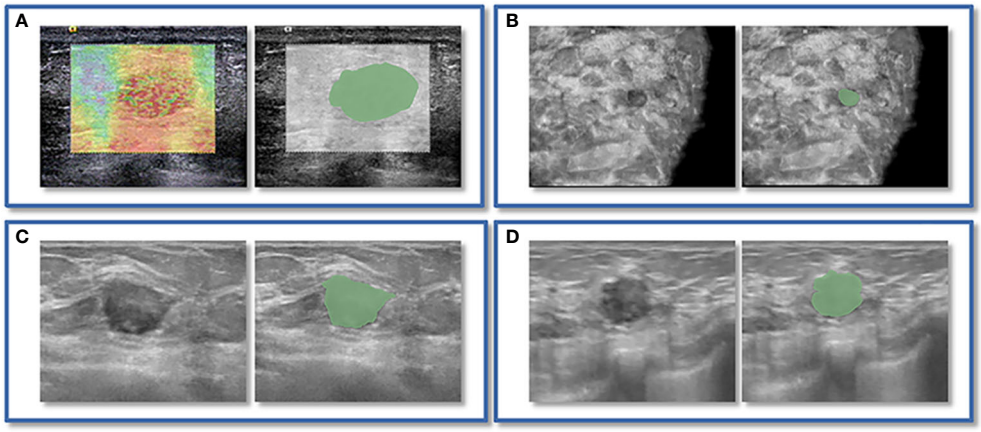


FIGURE 2
An instance of manually delineating a region of interest (ROI). The strain elastography (SE) and automated breast volume scanner (ABVS) images of a 41-year-old female with a solid hypoechoic lesion measuring approximately 16x11x12mm on her left breast. The lesion was irregular in shape, parallel in position, with still well-defined borders, sharp margins, and scattered microcalcifications visible internally, and exhibited no significant posterior echogenicity change or retraction in the coronal plane, and the ultrasound elasticity score was 4, finally, the lesion was classified as BI-RADS category 4a. Pathological examination confirmed it as invasive ductal carcinoma. ROI segmentation was performed on both the SE image **(A)** and ABVS coronal image **(B)**, with delineation along the boundary of the lesion followed by uniform outward expansion of its edges by 3 mm to encompass some surrounding tissue.ROI segmentation was performed on ABVS transverse **(C)** and sagittal **(D)** images, respectively, and meticulous delineation was performed along the lesion's contour and borders on these two views.

2.4 Development of models

2.4.1 Radiomics models

Based on radiomics features of SE and ABVS images, logistic regression analysis was utilized to build the radiomics model, including the SE radiomics model, the ABVS radiomics model, the SE+ABVS radiomics model that was developed by combining the two. The radiomics score (Radscore) for each lesion was computed by weighting the coefficients of features in the SE+ABVS radiomics model.

2.4.2 Clinical diagnostic model

The BI-RADS model was constructed through logistic regression analysis of lesion's BI-RADS categories in the training group.

2.5 Development and performance validation of nomogram

This study performed a univariate analysis in order to determine the risk predictive variables associated with breast cancer(P<0.05), which were then combined with the results of conventional imaging assessment and radiomics analysis. Based on these findings, we integrated relevant clinical risk factors, the Radscore, and the BI-RADS category of lesions to develop a nomogram for assessing the malignancy risk in such breast lesions by multivariate logistic regression analysis. Subsequently, the nomogram's diagnostic performance was compared to that of the BI-RADS model and SE +ABVS radiomics model. To evaluate the diagnostic performance of the models, we calculated the area under the receiver operating characteristic curve (AUC) for each model in the training group, validation group, and in the BI-RADS category 4 lesions within both groups. Furthermore, the DeLong test was used to examine differences in AUC values between different models. The nomogram's goodness of fit was investigated graphically and by calculating significance by plotting the calibration curve and conducting the Hosmer-Lemeshow test. Lastly, clinical decision analysis curves were drawn for quantifying the net benefits of the BI-RADS model, SE+ABVS radiomics models, and nomogram at various threshold probabilities.

2.6 Statistical analysis

SPSS 23.0, R 4.2.2, and MedCalc 19.6.0 were utilized for statistical analysis and graph plotting. The 'psych', 'survival', 'glmnet', 'rms', 'ResourceSelection', and 'rmda' packages were used in R. We performed normality tests on each group of data and selected the appropriate hypothesis test based on the results to compare the distribution of data between the training and validation groups. The study has chosen a significance level of 0.05 as the threshold for detecting statistical differences.

3 Results

3.1 Comparison of clinical basis information and sonographic features

The study included 423 breast lesions that were pathologically confirmed to include 215 benign lesions and 208 malignant lesions. Table 1 demonstrates that both the clinical basis data and sonographic features of lesions were evenly distributed in the training and validation groups, indicating no statistically significant differences between the two groups (P>0.05). Furthermore, the univariate risk analysis revealed that patients with malignant lesions had significantly higher age and lesion's maximum diameter compared to those with benign lesions in

10.3389/fonc.2023.1256146

TABLE 1 Clinical basis information and sonographic features of patients with breast lesions.

Characteristic	Training group (n=296)			Validation group (n=127)			
	Benign group (n=151)	Malignant group (n=145)	P_{Intra} value	Benign group (n=64)	Malignant group (n=63)	P_{Intra} *value*	P_{Inter} *value*
Age							
Median±SD	39±10.85	51±10.73	<0.01*	41±9.83	53±10.82	<0.01*	0.34
maximum diameter							
Median±SD	13±7.36	19±8.23	<0.01*	15±9.24	18±8.52	<0.05*	0.37
Location							
Left	72 (47.68%)	79 (54.48%)	0.24	32 (50%)	30 (47.62%)	0.79	0.68
right	79 (52.32%)	66 (45.52%)		32 (50%)	33 (52.39%)		
Morphology							
Regular	77 (50.99%)	17 (11.72%)	<0.01*	28 (43.75%)	5 (7.94%)	<0.01*	0.24
Irregular	74 (49.01%)	128 (88.28%)		36 (56.25%)	58 (92.06%)		
Border							
Clear	110 (72.85%)	45 (31.03%)	<0.01*	43 (67.19%)	19 (30.16%)	<0.01*	0.50
Not Clear	41 (27.15%)	100 (68.94%)		21 (32.81%)	44 (69.84%)		
Margin							
Circumscribed	127 (84.11%)	22 (15.17%)	<0.01*	51 (79.69%)	12 (19.05%)	<0.01*	0.89
Not circumscribed	24 (15.89%)	123 (84.83%)		13 (20.31%)	51 (80.95%)		
Orientation							
Parallel	128 (84.77%)	79 (54.48%)	<0.01*	54 (84.38%)	32 (50.79%)	<0.01*	0.65
Not parallel	23 (15.23%)	66 (45.52%)		10 (15.62%)	31 (49.21%)		
Posterior echogenicity							
Enhancement	11 (7.28%)	17 (11.72%)	0.28	4 (6.25%)	6 (9.52%)	0.82	0.52
No difference	130 (86.10%)	111 (76.56%)		57 (89.06%)	52 (82.54%)		
Shadowing	10 (6.62%)	17 (11.72%)		3 (4.69%)	5 (7.94%)		
Retraction sign							
Presence	0 (0%)	32 (22.07%)	<0.01*	0 (0%)	16 (25.40%)	<0.01*	0.60
Absence	151 (100%)	113 (77.93%)		64 (100%)	47 (74.60%)		
MicroCalcification							
Presence	16 (10.60%)	66 (45.52%)	<0.01*	3 (4.69%)	26 (41.27%)	<0.01*	0.30
Absence	135 (89.40%)	79 (54.48%)		61 (95.31%)	37 (58.73%)		
Ultrasonic elasticity score							
1 point	16 (10.60%)	0 (0%)		7 (10.94%)	0 (0%)		
2 points	49 (32.45%)	1 (0.69%)		22 (34.37%)	3 (4.76%)		
3 points	65 (43.05%)	21 (14.48%)	<0.01*	27 (42.19%)	14 (22.22%)	<0.01*	0.70
4 points	20 (13.24%)	55 (37.93%)		8 (12.50%)	24 (38.10%)		
5 points	1 (0.66%)	68 (46.90%)		0 (0%)	22 (34.92%)		

(Continued)

TABLE 1 Continued

Characteristic	Training group (n=296)			Validation group (n=127)			
	Benign group (n=151)	Malignant group (n=145)	P_{Intra} value	Benign group (n=64)	Malignant group (n=63)	P_{Intra} *value*	P_{Inter} *value*
BI-RADS							
3	64 (42.38%)	4 (2.76%)		27 (42.19%)	2 (3.17%)		
4a	82 (54.31%)	21 (14.48%)	<0.01*	34 (53.12%)	11 (17.46%)	<0.01*	0.77
4b	4 (2.65%)	27 (18.62%)		3 (4.69%)	12 (19.05%)		
4c	1 (0.66%)	30 (20.69%)		0 (0%)	17 (26.98%)		
5	0 (0%)	63 (43.45%)		0 (0%)	21 (33.34%)		

* $p < 0.05$.

both groups (P<0.05). However, no correlation was seen between the location of the lesion and the malignant risk of the lesion (P>0.05). Hence, we regarded age and lesion size as predictor variables in the context of breast cancer. Regarding the sonographic features of the lesions, there were statistically significant differences (P<0.01) observed in morphology, borders, margins, orientation, microcalcifications, retraction condition of the coronal plane, and elasticity scores between benign and malignant lesions within both groups. while no statistical differences were found in posterior echogenicity (P>0.05). This study assessed the malignancy risk of lesions by these sonographic features of them, and the BI-RADS categories obtained were also significantly different in benign and malignant lesions (P<0.01).

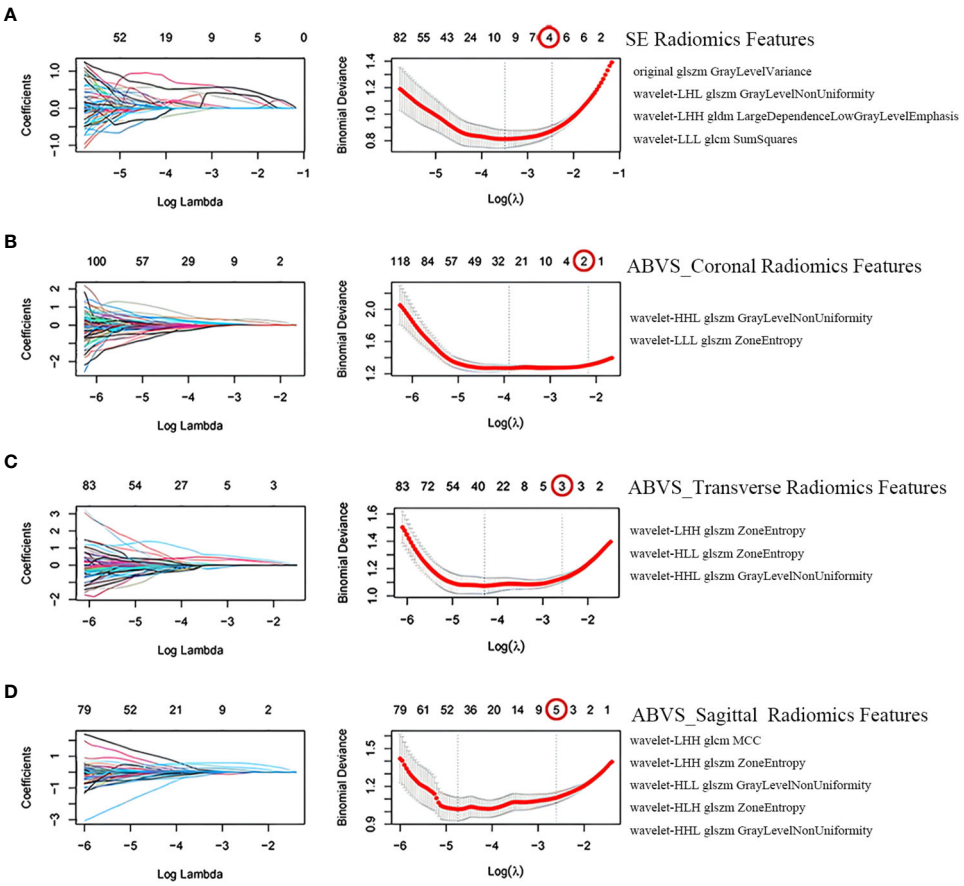


FIGURE 3
Screening of radiomics features. Selection of strain elastography (SE) radiomics features **(A)**, automated breast volume scanner (ABVS) coronal plane radiomics features **(B)**, ABVS transverse plane radiomics features **(C)**, and ABVS sagittal plane radiomics features **(D)** using the least absolute shrinkage and selection operator (LASSO) regression model. The coefficient profiles of LASSO for each modal radiomics feature are presented on the left. The right shows that the tuning parameter λ (lambda) in the LASSO model was selected using tenfold cross-validation, and the binomial deviance was plotted as a function of log(λ), with vertical dashed lines drawn at the minimum deviation (log(λ.min)) and the 1 standard error of the minimum deviation (log(λ.1se)). Selected the non-zero coefficient features in the model when the horizontal coordinate was log(λ.1se).

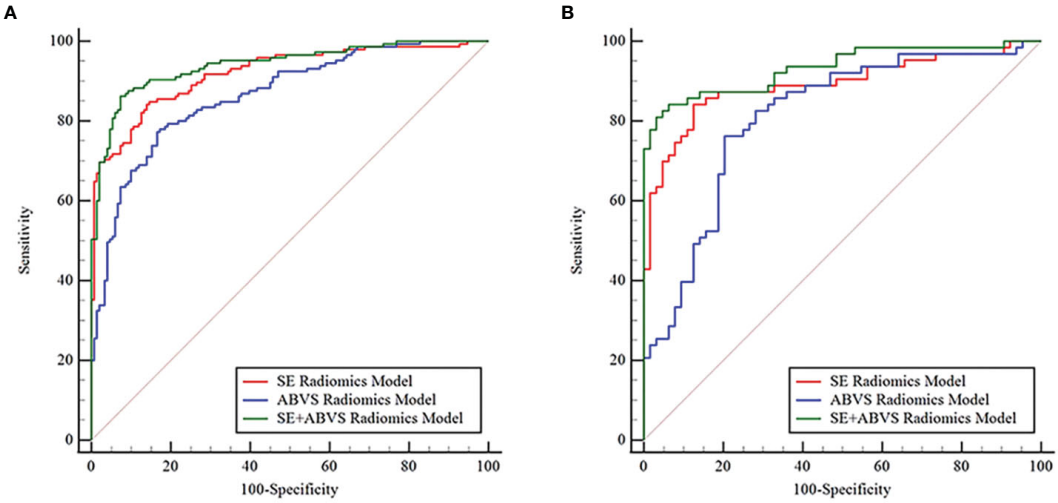


FIGURE 4
The receiver operator characteristic curves for various radiomics models in the training **(A)** and validation groups **(B)**.

3.2.Screening of radiomics features

The SE and ABVS images generated 837 and 2511 radiomics features (ABVS cross plane, sagittal plane, and coronal plane, each generated 837 features), respectively. The training group's SE radiomics features as well as the ABVS coronal plane, transverse plane, and sagittal plane radiomics features underwent sequential ICC analysis, Mann-Whitney U test, LASSO regression analysis with tenfold cross-validation for dimensionality reduction. Finally, a total of 14 features were selected, comprising four SE radiomics features and ten ABVS radiomics features (two from the coronal plane, three from the transverse plane, and five from the sagittal plane). All of these radiomics features are texture features, one of which was from the original image and thirteen were obtained after wavelet transform (Figure 3).

3.3 Comparison of radiomics models

By comparing and validating the diagnostic efficacy of the radiomics models (Figure 4, Table 2), the AUC values of the selected ABVS and SE features for distinguishing between benign and malignant solid hypoechoic breast lesions were consistently above 0.8 in both the training and validation groups. Moreover, compared to any single-modality radiomics models, the SE+ABVS radiomics model, which integrated the radiomics features of two imaging modalities, demonstrated significantly higher AUC values in both training (All $P<0.01$) and validation groups (compared to the ABVS radiomics model: $P<0.01$, compared to SE radiomics model: $P<0.05$). These outcomes suggest that combining radiomics features from both SE and ABVS could enhance the accuracy of diagnostic models. Thus, the Radscore for each patient was obtained by weighting the corresponding coefficients for each feature in the SE+ABVS radiomics model., the formula is shown below, the Radscore for malignant lesions was found to be significantly higher than that for benign lesions within both groups. (Training group: 2.86 + 2.66, -2.34 + 1.80, $P<0.01$; Validation group: 2.60 + 2.31, -2.17 + 1.85, $P<0.01$).

TABLE 2 The AUC values of radiomics models in the training and validation groups.

Model	AUC (95%CI)	*P* (AUC compare to SE radiomics model)	*P* (AUC compare to ABVS radiomics model)
Training group			
SE radiomics model	0.920 (0.883, 0.948)		<0.01*
ABVS radiomics model	0.865 (0.821, 0.902)	<0.01*	
SE+ABVS radiomics model	0.941 (0.907, 0.965)	<0.01*	<0.01*
Validation group			
SE radiomics model	0.892 (0.824, 0.940)		0.08
ABVS radiomics model	0.811 (0.732,0.875)	0.08	
SE+ABVS radiomics model	0.933 (0.875,0.970)	<0.05*	<0.01*

*P< 0.05.

Radscore = 1.944469*original glszm GrayLevelVariance _ SE
+0.03492*wavelet – LHL glszm GrayLevelNonUniformity _ SE
+0.634491*wavelet – LHH gldm LargeDependenceLowGrayLevelEmphasis _ SE
+0.465968*wavelet – LLL glcm SumSquares _ SE
+0.133391*wavelet – HHL glszm GrayLevelNonUniformity _ ABVS _ Coronal
−1.515639*wavelet – LLL glszm ZoneEntropy _ ABVS _ Coronal
+0.010666*wavelet – LHH glszm ZoneEntropy _ ABVS _ Transverse
+1.292836*wavelet – HLL glszm ZoneEntropy _ ABVS _ Transverse
+0.043941*wavelet – HHL glszm GrayLevelNonUniformity _ ABVS _ Transverse
−1.828919*wavelet – LHH glcm MCC _ ABVS _ Sagittal
+0.813141*wavelet – LHH glszm ZoneEntropy _ ABVS _ Sagittal
−0.008279*wavelet – HLL glszm GrayLevelNonUniformity _ ABVS _ Sagittal
+1.520859*wavelet – HLH glszm ZoneEntropy _ ABVS _ Sagittal
+0.271946*wavelet – HHL glszm GrayLevelNonUniformity _ ABVS _ Sagittal
−25.66532

3.4 Evaluation of nomogram performance

Based on the clinical risk factors identified through univariate analysis, BI-RADS categories determined from imaging assessments, and Radscore obtained from radiomics analysis, we constructed a nomogram using multivariate logistic regression to visually assess the risk of malignancy in solid hypoechoic breast lesions, the nomogram incorporated the patient's age, lesion's maximum diameter, Radscore, and BI-RADS category. As illustrated in Figure 5, Radscore had the highest weightage followed by BI-RADS score while age and maximum diameter of the lesion exerted less influence on assessment results.

Figure 6 and Table 3 present that the BI-RADS model, SE +ABVS radiomics model, and nomogram are effective in predicting the malignancy risk in solid hypoechoic breast lesions, Notably, the nomogram exhibits superior diagnostic performance with higher AUC values (0.972, 0.964) in training and validation group compared to both the BI-RADS model (AUC: 0.930, 0.916) and SE+ABVS radiomics models (AUC: 0.941, 0.933). Furthermore, its difference with BI-RADS model and SE+ABVS radiomics model was statistically significant in both groups (P<0.05). Besides, we further compared the diagnostic efficacy of the three models for BI-RADS category 4 lesions within the two groups. The results revealed that the nomogram (AUC: 0.952, 0.930) consistently exhibited higher AUC values than both the BI-RADS model (AUC:0.844, 0.839) and SE+ABVS radiomics model (AUC:0.915, 0.899). Moreover, there were consistently statistically significant differences between the nomogram and BI-RADS model (All P<0.01). However, in comparison to the SE+ABVS model, the nomogram was only statistically different from it in the training group (P<0.05), but not in the validation group (P>0.05). Other than that, in terms of diagnostic sensitivity, specificity, and accuracy, Although the specificity of the nomogram was slightly inferior to that of the BI-RADS model in the training group, it significantly improved diagnostic sensitivity. Furthermore, its diagnostic parameters were at the highest level across all validation groups. These results suggest that the nomogram exhibited the best overall diagnostic performance. Finally, we observed that the AUC values of the SE+ABVS Radiomics model consistently outperformed those of the BI-RADS model, and a statistically significant difference was found between them when diagnosing BI-RADS category 4 lesions of the training group (P<0.05). This finding highlights the ability of radiomics analysis to detect deep-seated features within the images, ultimately leading to improved diagnostic efficiency.

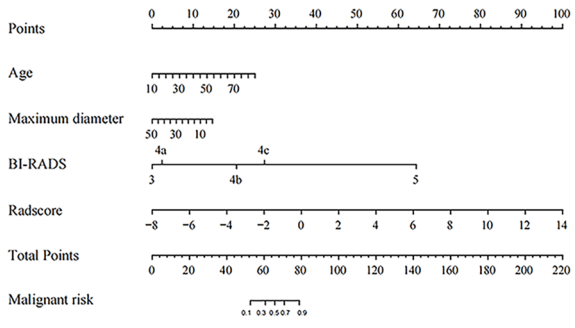


FIGURE 5
The Nomogram for predicting the malignant risk of solid hypoechoic breast lesions.

The calibration curve exhibits a favorable fit of the nomogram (Figures 7A, B). indicating that the predicted risk by the nomogram was close to the observed risks. The results from the Hosmer-Lemeshow test further proved that the differences between them did not present statistical significance in either the training group (P=0.70) or validation group (P=0.95).

The clinical decision analysis curve (Figure 7C) indicates that utilizing the BI-RADS model, SE+ABVS radiomics model, and nomogram for decision-making significantly improved the net benefit for patients compared to the assumption of intervention for all lesions or no intervention at all. Furthermore, the nomogram provided a greater net benefit to patients compared to both the BI-RADS model and SE+ABVS radiomics model.

4 Discussion

The study combined radiomics features of ABVS and SE images with conventional imaging diagnosis criteria along with clinical risk factors for developing a clinical-radiomics nomogram that demonstrated excellent diagnostic efficacy, as well as good calibration capabilities, and significant clinical usefulness.

Although ABVS and SE examination techniques offer significant advantages in breast screening, the examiner's naked eye remains incapable of capturing deep image information. Radiomics provides a pathway to capture internal tumor information at a more profound level. Wang et al. derived radiomics features from ABVS images and constructed multiple machine learning models for breast cancer diagnosis, the best of which was the support vector machine model with an AUC of 0.857 (17). Additionally, Liu et al. employed radiomics features extracted from SE images for breast cancer prediction, yielding a Radscore with an AUC of 0.866 in the test set (18). Besides, Ma et al. developed a multivariate logistic

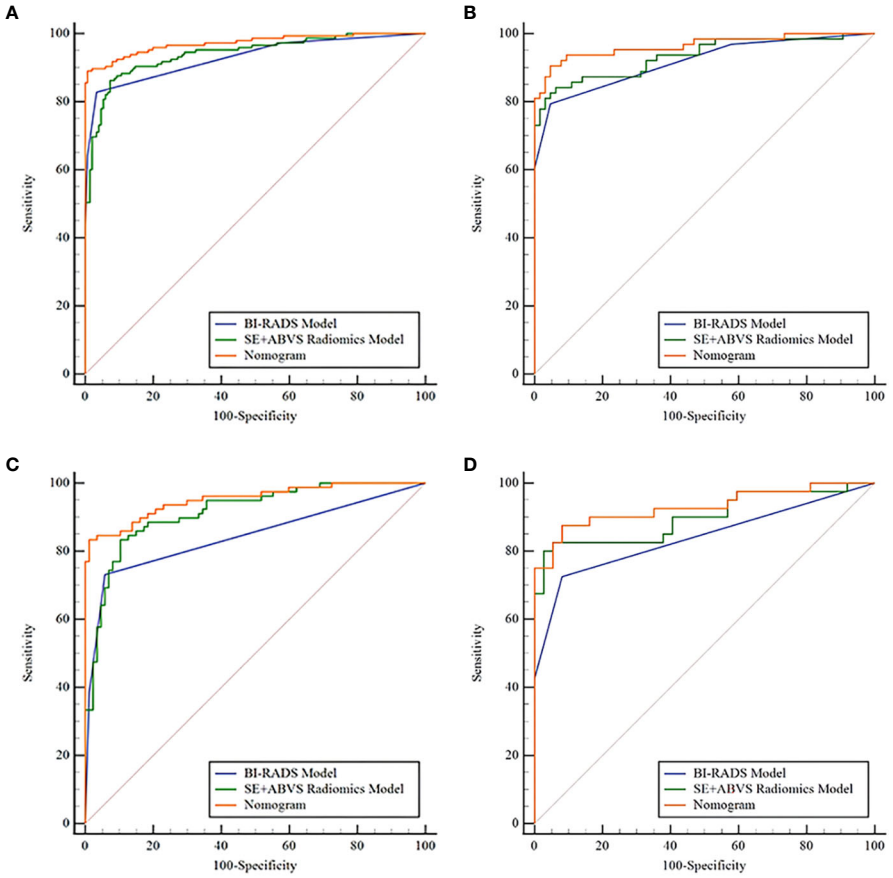


FIGURE 6
The receiver operator characteristic curves of the BI-RADS model, SE+ABVS radiomics model, and Nomogram in the training group **(A)**, the validation group **(B)**, the BI-RADS category 4 lesions in the training group **(C)**, and the BI-RADS category 4 lesions in the validation group **(D)**.

model by combining SE, B-mode, and ABVS coronal radiomics features, with an AUC value of 0.946 in the internal validation group (19). In this study, we performed radiomics analysis on both ABVS and SE images. For ABVS images, we delineated the ROI across sagittal, transverse, and coronal planes. While outlining the ROI in the ABVS coronal planes and the SE image, we incorporated a portion of the lesion's peripheral tissues to capture additional information. As a result, the ABVS and SE features that we acquired demonstrated good predictive capabilities for breast cancer, and the combination of the two yielded a higher diagnostic efficacy than the BI-RADS model that obtained by a highly experienced physician based on visual assessment alone (AUC: 0.933 vs. 0.916). Additionally, this study analyzed clinical risk factors related to breast cancer and revealed that age and lesion size exhibited significantly higher values in the malignant group compared to the benign group, which is consistent with previous research findings (33–35). Therefore, we developed a nomogram by integrating Radscore, patient's age, maximum diameter of the lesion, and BI-RADS scores using multivariate logistic regression analysis. The AUC of this nomogram in the internal validation group was 0.964, which surpassed that of both the SE+ABVS radiomics model and the clinical model. Furthermore, we conducted an analysis on the clinical utility of this nomogram, and the decision analysis curves revealed that it could offer superior net benefit to patients across a broad range of threshold intervals. Consequently, the nomogram holds significant value as a point of reference for clinicians, particularly novice practitioners lacking diagnostic expertise in identifying suspicious lesions.

In addition, this nomogram has demonstrated significant advantages in the diagnosis of BI-RADS category 4 lesions. The appearance of these lesions on imaging can be highly deceptive, so they span a wide range of malignancy risks (36, 37), which makes clinical diagnosis extremely challenging, often necessitating biopsies to definitively determine the nature of such lesions (30). However, routine biopsy results are often influenced by the spatial heterogeneity of the lesion and operator expertise (38), while also being an invasive procedure with potential complications such as bleeding (39). The majority of radiomics studies for this category of lesions have predominantly utilized MR images, and these studies have yielded favorable outcomes (26–29). Nevertheless, MR examinations are expensive, time-consuming, and not suitable for common screenings (40, 41). Based on ABVS images, Wang et al. integrated clinical ultrasound factors and Radscore to develop a nomogram for the diagnosis of BI-RADS category 4 lesions, which achieved an AUC value of 0.925 in the internal validation group and effectively minimized unnecessary biopsies (20). During this study, we constructed nomogram that also achieved an AUC value of 0.930 for the diagnosis of BI-RADS category 4 lesions in the

TABLE 3 The diagnostic parameters of the BI-RADS model, SE+ABVS radiomics model, and Nomogram in each group.

Model	AUC (95%CI)	Sensitivity %	Specificity %	Accuracy %	*P* (AUC compare to BI-RADS model)	*P* (AUC compare to SE+ABVS radiomics model)
Training group						
BI-RADS model	0.930(0.894, 0.956)	82.76	96.69	89.86		0.46
SE+ABVS radiomics model	0.941(0.907, 0.965)	87.59	90.07	88.85	0.46.	
Nomogram	0.972(0.946, 0.988)	89.66	94.04	91.89	<0.01*	<0.01*
Validation group						
BI-RADS model	0.916(0.853, 0.958)	79.37	95.31	87.40		0.52
SE+ABVS radiomics model	0.933(0.875, 0.970)	85.71	85.94	85.83	0.52	
Nomogram	0.964(0.916, 0.989)	87.30	95.31	91.34	<0.05*	<0.05*
BI-RADS category 4 lesions in the Training group						
BI-RADS model	0.844(0.779, 0.895)	73.08	94.25	84.24		<0.05*
SE+ABVS radiomics model	0.915(0.862, 0.953)	84.62	87.36	86.06	<0.05*	
Nomogram	0.952(0.907, 0.979)	84.62	89.66	87.27	<0.01*	<0.05*
BI-RADS category 4 lesions in the Validation group						
BI-RADS model	0.839(0.738, 0.913)	72.50	91.89	81.82		0.13
SE+ABVS radiomics model	0.899(0.809, 0.956)	82.50	83.78	83.12	0.13	
Nomogram	0.930(0.848, 0.975)	82.50	91.89	87.01	<0.01*	0.12

*P<0.05.

validation group, surpassing the performance of the clinical model (AUC: 0.839), thereby further validating its good diagnostic efficacy. This may be attributed to the fact that the radiomics features selected for this study are all texture features with the majority derived from wavelet transform. Previous studies have demonstrated the value of wavelet transform-based texture features for the diagnosis of tumor lesions (42). The primary advantage of wavelet transform in image analysis lies in its multi-scale analysis capability, allowing it to capture the texture information of an image at various granularities. It possesses directional sensitivity, enabling it to accurately identify texture changes in multiple directions, while its time-frequency localization property allows it to keenly detect local variations in images. Additionally, wavelet transform can enhance image contrast, exhibit certain resistance to noise, and effectively compress image information, making feature extraction more robust and efficient (43). By quantifying the textural variances of breast lesions, we successfully captured the subtle heterogeneity within these lesions, thereby effectively distinguished between benign and malignant breast lesions.

Lambin et al. introduced the radiomics quality score (RQS) to provide a framework for clinical researchers to evaluate and guide their radiomics studies (44). This study has given comparatively detailed elaboration on image acquisition, feature extraction and screening, and model construction in order to ensure the reproducibility of the study. Two physicians independently delineated the lesions, effectively achieving multiple segmentations. The features extracted from both segmentations were then subjected to ICC analysis. Consequently, only the features demonstrating excellent repeatability and robustness were selected. To prevent model overfitting, we standardized the feature values. Features with strong discriminative ability were obtained through the U-test, LASSO regression with tenfold cross-validation

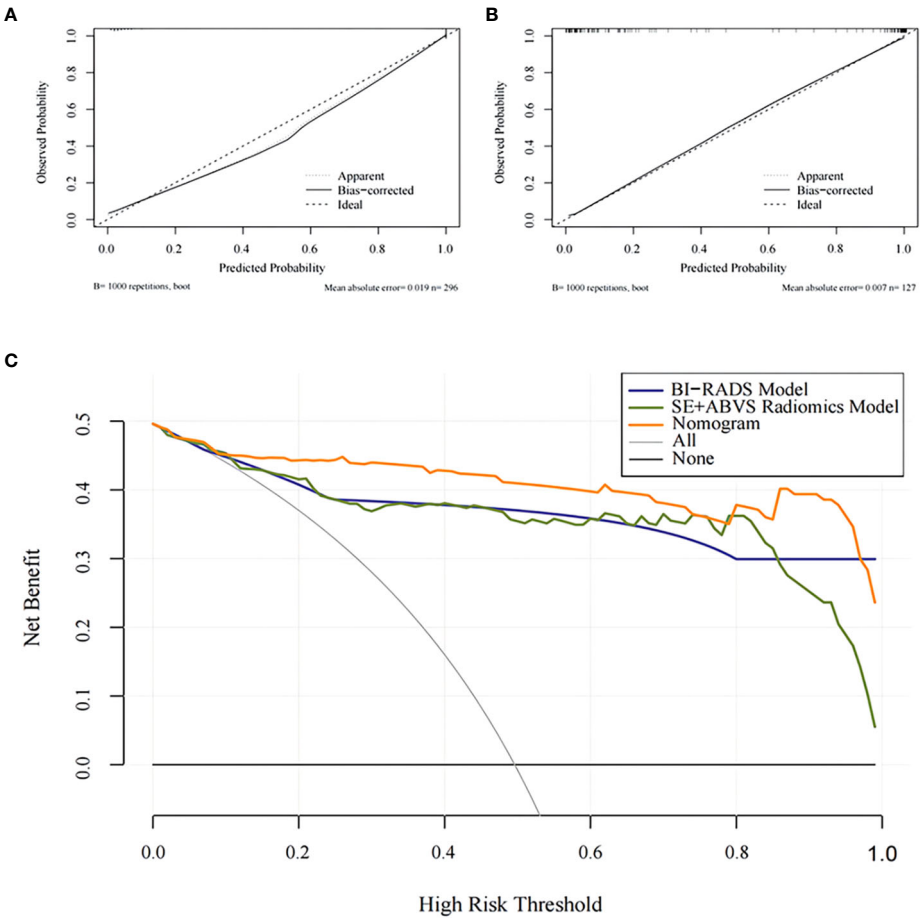


FIGURE 7
The calibration curves for the Nomogram in the training group **(A)** and the validation group **(B)**.The decision analysis curves of the BI-RADS mode, SE+ABVS radiomics model, and Nomogram in the validation group **(C)**.

from ABVS and SE images, respectively. Subsequently, we evaluated the constructed nomogram using calibration curves and the Hosmer-Lemeshow test, which demonstrated excellent calibration performance. Based on these analyses, the nomogram appears to be a robust and generalizable tool, offering accurate risk prediction with potential for practical clinical implementation. Although the development of this nomogram necessitates a combination of diverse factors, these data can be retrospectively obtained without imposing an additional examination burden on patients.

Admittedly, This study was subject to certain limitations: It was conducted as a single-center retrospectively study thus selection bias may have occurred, and lacked external validation, which necessitates further multicenter large-sample studies and prospective trials for the validation of our developed nomogram.

5 Conclusion

The nomogram developed in this study, which combined SE and ABVS radiomics features, with traditional imaging assessment criteria and clinical risk factors, it can serve as a reliable and non-invasive analytical tool to assist physicians in accurately assessing the malignancy risk in solid hypoechoic breast lesions, leading to better clinical decision-making.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Institutional Review Board of the Third Xiangya Hospital, CSU. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

GS: Conceptualization, Formal Analysis, Methodology, Writing – original draft, Writing – review & editing. JL: Formal Analysis, Methodology, Writing – review & editing. YY: Data curation, Writing – review & editing. ZY: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Supervision, Writing – review & editing.

Funding

This work was supported by grants from the Natural Science Foundation of Hunan Province, China (No. 2022JJ40760) and Project of Hunan Provincial Health Commission (No.202209024823).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. AlSendi M, O'Reilly D, Zeidan YH, Kelly CM. Oligometastatic breast cancer: Are we there yet? *Int J Cancer* (2021) 149(8):1520–8. doi: 10.1002/ijc.33693

3. Liang Y, Zhang H, Song X, Yang Q. Metastatic heterogeneity of breast cancer: Molecular mechanism and potential therapeutic targets. *Semin Cancer Biol* (2020) 60:14–27. doi: 10.1016/j.semcancer.2019.08.012

4. Cedolini C, Bertozzi S, Londero AP, Bernardi S, Seriau L, Concina S, et al. Type of breast cancer diagnosis, screening, and survival. *Clin Breast Cancer* (2014) 14(4):235–40. doi: 10.1016/j.clbc.2014.02.004

5. Yang R, Han Y, Yi W, Long Q. Autoantibodies as biomarkers for breast cancer diagnosis and prognosis. *Front Immunol* (2022) 13:1035402. doi: 10.3389/fimmu.2022.1035402

6. Zhi H, Ou B, Luo BM, Feng X, Wen YL, Yang HY. Comparison of ultrasound elastography, mammography, and sonography in the diagnosis of solid breast lesions. *J ultrasound Med* (2007) 26(6):807–15. doi: 10.7863/jum.2007.26.6.807

7. Balleyguier C, Ciolovan L, Ammari S, et al. Breast elastography: the technical process and its applications. *Diagn interventional Imaging* (2013) 94(5):503–13. doi: 10.1016/j.diii.2013.02.006

8. Golatta M, Franz D, Harcos A, Junkermann H, Rauch G, Scharf A, et al. Interobserver reliability of automated breast volume scanner (ABVS) interpretation and agreement of ABVS findings with hand held breast ultrasound (HHUS), mammography and pathology results. *Eur J Radiol* (2013) 82(8):e332–6. doi: 10.1016/j.ejrad.2013.03.005

9. Chen L, Chen Y, Diao XH, Fang L, Pang Y, Cheng AQ, Fang L, Pang Y, Cheng AQ, et al. Comparative study of automated breast 3-D ultrasound and handheld B-mode ultrasound for differentiation of benign and Malignant breast masses. *Ultrasound Med Biol* (2013) 39(10):1735–42. doi: 10.1016/j.ultrasmedbio.2013.04.003

10. Zheng FY, Yan LX, Huang BJ, Xia HS, Wang X, Lu Q, et al. Comparison of retraction phenomenon and BI-RADS-US descriptors in differentiating benign and Malignant breast masses using an automated breast volume scanner. *Eur J Radiol* (2015) 84(11):2123–9. doi: 10.1016/j.ejrad.2015.07.028

11. Xu C, Wei S, Xie Y, Guan X, Fu N, Huang P, et al. Combined use of the automated breast volume scanner and the US elastography for the differentiation of benign from Malignant lesions of the breast. *BMC Cancer* (2014) 14:798. doi: 10.1186/1471-2407-14-798

12. Guo R, Lu G, Qin B, Fei B. Ultrasound imaging technologies for breast cancer detection and management: A review. *Ultrasound Med Biol* (2018) 44(1):37–70. doi: 10.1016/j.ultrasmedbio.2017.09.012

13. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169

14. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer (Oxford England: 1990)* (2012) 48(4):441–6. doi: 10.1016/j.ejca.2011.11.036

15. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol* (2017) 72(1):3–10. doi: 10.1016/j.crad.2016.09.013

16. Luo WQ, Huang QX, Huang XW, Hu HT, Zeng FQ, Wang W. Predicting breast cancer in breast imaging reporting and data system (BI-RADS) ultrasound category 4 or 5 lesions: A nomogram combining radiomics and BI-RADS. *Sci Rep* (2019) 9(1):11921. doi: 10.1038/s41598-019-48488-4

17. Wang H, Yang X, Ma S, Zhu K, Guo S. An optimized radiomics model based on automated breast volume scan images to identify breast lesions: comparison of machine learning methods: comparison of machine learning methods. *J Ultrasound Med* (2022) 41(7):1643–55. doi: 10.1002/jum.15845

18. Liu X, Zhang J, Zhou J, He Y, Xu Y, Zhang Z, et al. Multi-modality radiomics nomogram based on DCE-MRI and ultrasound images for benign and Malignant breast lesion classification. *Front Oncol* (2022) 12:992509. doi: 10.3389/fonc.2022.992509

19. Ma Q, Shen C, Gao Y, Duan Y, Li W, Lu G, et al. Radiomics analysis of breast lesions in combination with coronal plane of ABVS and strain elastography. *Breast Cancer (Dove Med Press)* (2023) 15:381–90. doi: 10.2147/BCTT.S410356

20. Wang SJ, HQ L, Yang T, Huang MQ, Zheng BW, Wu T, et al. Automated breast volume scanner (ABVS)-based radiomic nomogram: A potential tool for reducing unnecessary biopsies of BI-RADS 4 lesions. *Diagnostics (Basel)* (2022) 12(1):172. doi: 10.3390/diagnostics12010172

21. Lin F, Wang Z, Zhang K, Yang P, Ma H, Shi Y, et al. Contrast-enhanced spectral mammography-based radiomics nomogram for identifying benign and Malignant breast lesions of sub-1 cm. *Front Oncol* (2020) 10:573630. doi: 10.3389/fonc.2020.573630

22. Parekh VS, Jacobs MA. Integrated radiomic framework for breast cancer and tumor biology using advanced machine learning and multiparametric MRI. *NPJ Breast Cancer* (2017) 3:43. doi: 10.1038/s41523-017-0045-3

23. Zhang Q, Xiao Y, Suo J, Shi J, Yu J, Guo Y, et al. Sonoelastomics for breast tumor classification: A radiomics approach with clustering-based feature selection on sonoelastography. *Ultrasound Med Biol* (2017) 43(5):1058–69. doi: 10.1016/j.ultrasmedbio.2016.12.01

24. Xu Z, Wang Y, Chen M, Zhang Q. Multi-region radiomics for artificially intelligent diagnosis of breast cancer using multimodal ultrasound. *Comput Biol Med* (2022) 149:105920. doi: 10.1016/j.compbiomed.2022.105920

25. Ou X, Wang J, Zhou R, Zhu S, Pang F, Zhou Y, et al. Ability of F-FDG PET/CT radiomic features to distinguish breast carcinoma from breast lymphoma. *Contrast Media Mol Imaging* (2019) 2019:4507694. doi: 10.1155/2019/4507694

26. Zhang R, Wei W, Li R, Li J, Zhou Z, Ma M, et al. An MRI-based radiomics model for predicting the benignity and Malignancy of BI-RADS 4 breast lesions. *Front Oncol* (2022) 11:733260. doi: 10.3389/fonc.2021.733260

27. Cui Q, Sun L, Zhang Y, Zhao Z, Li S, Liu Y, et al. Value of breast MRI omics features and clinical characteristics in Breast Imaging Reporting and Data System (BI-RADS) category 4 breast lesions: an analysis of radiomics-based diagnosis. *Ann Transl Med* (2021) 9(22):1677. doi: 10.21037/atm-21-5441

28. Lyu Y, Chen Y, Meng L, Guo J, Zhan X, Chen Z, et al. Combination of ultrafast dynamic contrast-enhanced MRI-based radiomics and artificial neural network in assessing BI-RADS 4 breast lesions: Potential to avoid unnecessary biopsies. *Front Oncol* (2023) 13:1074060. doi: 10.3389/fonc.2023.1074060

29. Hao W, Gong J, Wang S, Zhu H, Zhao B, Peng W. Application of MRI radiomics-based machine learning model to improve contralateral BI-RADS 4 lesion assessment. *Front Oncol* (2020) 10:531476. doi: 10.3389/fonc.2020.531476

30. Mendelson EB, Böhm-Vélez M, Berg WA, et al. ACR BIRADS® Ultrasound, in: *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. American College of

Radiology, Reston, 2013 VA. Available at: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-rads (Accessed 11 Jan 2023).

31. Zhi H, Xiao XY, Ou B, Zhong WJ, Zhao ZZ, Zhao XB, et al. Could ultrasonic elastography help the diagnosis of small (≤2 cm) breast cancer with the usage of sonographic BI-RADS classification? *Eur J Radiol* (2012) 81(11):3216–21. doi: 10.1016/j.ejrad.2012.04.016

32. Büsing KA, Kilian AK, Schaible T, Debus A, Weiss C, Neff KW. Reliability and validity of MR image lung volume measurement in fetuses with congenital diaphragmatic hernia and in *vitro* lung models. *Radiology* (2008) 246(2):553–61. doi: 10.1148/radiol.2462062166

33. Marzbani B, Nazari J, Najafi F, Marzbani B, Shahabadi S, Amini M, et al. Dietary patterns, nutrition, and risk of breast cancer: a case-control study in the west of Iran. *Epidemiol Health* (2019) 41:e2019003. doi: 10.4178/epih.e2019003

34. Sun YS, Zhao Z, Yang ZN, Xu F, Lu HJ, Zhu ZY, et al. Risk factors and preventions of breast cancer. *Int J Biol Sci* (2017) 13(11):1387–97. doi: 10.7150/ijbs.21635

35. Gity M, Arabkheradmand A, Taheri E, Shakiba M. Diagnostic investigation of breast magnetic resonance imaging in Malignant and benign mass lesions. *Arch Med Sci* (2018) 14(5):1061–9. doi: 10.5114/aoms.2016.62281

36. Zou X, Wang J, Lan X, Lin Q, Han F, Liu L, Lin Q, Han F, Liu L, et al. Assessment of diagnostic accuracy and efficiency of categories 4 and 5 of the second edition of the BI-RADS ultrasound lexicon in diagnosing breast lesions. *Ultrasound Med Biol* (2016) 42(9):2065–71. doi: 10.1016/j.ultrasmedbio.2016.04.020

37. Liu G, Zhang MK, He Y, Liu Y, Li XR, Wang ZL. BI-RADS 4 breast lesions: could multi-mode ultrasound be helpful for their diagnosis? *Gland Surg* (2019) 8(3):258–70. doi: 10.21037/gs.2019.05.01

38. Teng R, Wei Q, Zhou J, Dong M, Jin L, Hu W, et al. The influence of preoperative biopsy on the surgical method in breast cancer patients: a single-center experience of 3,966 cases in China. *Gland Surg* (2021) 10(3):1038–45. doi: 10.21037/gs-21-7

39. Nakano S, Imawari Y, Mibu A, Otsuka M, Oinuma T. Differentiating vacuum-assisted breast biopsy from core needle biopsy: Is it necessary? *Br J Radiol* (2018) 91 (1092):20180250. doi: 10.1259/bjr.20180250

40. Wang L. Early diagnosis of breast cancer. *Sensors (Basel)* (2017) 17(7):1572. doi: 10.3390/s17071572

41. Raber B, VJ B, Bedrosian I. How does MR imaging help care for my breast cancer patient? Perspective of a surgical oncologist. *Magn Reson Imaging Clin N Am* (2018) 26 (2):281–8. doi: 10.1016/j.mric.2017.12.010

42. Jiang Z, Yin J, Han P, Chen N, Kang Q, Qiu Y, et al. Wavelet transformation can enhance computed tomography texture features: a multicenter radiomics study for grade assessment of COVID-19 pulmonary lesions. *Quant Imaging Med Surg* (2022) 12 (10):4758–70. doi: 10.21037/qims-22-252

43. Sudarshan VK, Mookiah MR, Acharya UR, Chandran V, Molinari F, Fujita H, et al. Application of wavelet techniques for cancer diagnosis using ultrasound images: A Review. *Comput Biol Med* (2016) 69:97–111. doi: 10.1016/j.compbiomed.2015.12.006

44. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* (2017) 14(12):749–62. doi: 10.1038/nrclinonc.2017.141

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 03 November 2023
DOI 10.3389/fonc.2023.1274557

OPEN ACCESS

EDITED BY
Luca Nicosia,
European Institute of Oncology (IEO), Italy

REVIEWED BY
Ujjwal Agarwal,
Tata Memorial Hospital, India
Mehmet Dalmis,
Meta, Netherlands

*CORRESPONDENCE
Zhemin Zhuang
zmzhuang@stu.edu.cn
Shuxin Zhuang
zhuangshx3@mail2.sysu.edu.cn

†PRESENT ADDRESS
Shuxin Zhuang,
School of Biomedical Engineering,
Sun Yat-sen University, Shenzhen,
Guangdong, China

RECEIVED 10 August 2023
ACCEPTED 16 October 2023
PUBLISHED 03 November 2023

CITATION
Qiu S, Zhuang S, Li B, Wang J and Zhuang Z (2023) Prospective assessment of breast lesions AI classification model based on ultrasound dynamic videos and ACR BI-RADS characteristics.
Front. Oncol. 13:1274557.
doi: 10.3389/fonc.2023.1274557

Prospective assessment of breast lesions AI classification model based on ultrasound dynamic videos and ACR BI-RADS characteristics

Shunmin Qiu[1], Shuxin Zhuang[2*†], Bin Li[3], Jinhong Wang[4] and Zhemin Zhuang[5*]

[1]Department of Ultrasound, First Affiliated Hospital of Shantou University Medical College, Shantou, Guangdong, China, [2]School of Biomedical Engineering, Sun Yat-sen University, Shenzhen, Guangdong, China, [3]Product Development Department, Shantou Institute of Ultrasonic Instruments, Shantou, Guangdong, China, [4]Department of Ultrasound, Shantou Chaonan Minsheng Hospital, Shantou, Guangdong, China, [5]Engineering College, Shantou University, Shantou, Guangdong, China

Introduction: AI-assisted ultrasound diagnosis is considered a fast and accurate new method that can reduce the subjective and experience-dependent nature of handheld ultrasound. In order to meet clinical diagnostic needs better, we first proposed a breast lesions AI classification model based on ultrasound dynamic videos and ACR BI-RADS characteristics (hereafter, Auto BI-RADS). In this study, we prospectively verify its performance.

Methods: In this study, the model development was based on retrospective data including 480 ultrasound dynamic videos equivalent to 18122 static images of pathologically proven breast lesions from 420 patients. A total of 292 breast lesions ultrasound dynamic videos from the internal and external hospital were prospectively tested by Auto BI-RADS. The performance of Auto BI-RADS was compared with both experienced and junior radiologists using the DeLong method, Kappa test, and McNemar test.

Results: The Auto BI-RADS achieved an accuracy, sensitivity, and specificity of 0.87, 0.93, and 0.81, respectively. The consistency of the BI-RADS category between Auto BI-RADS and the experienced group (Kappa:0.82) was higher than that of the juniors (Kappa:0.60). The consistency rates between Auto BI-RADS and the experienced group were higher than those between Auto BI-RADS and the junior group for shape (93% vs. 80%; P = .01), orientation (90% vs. 84%; P = .02), margin (84% vs. 71%; P = .01), echo pattern (69% vs. 56%; P = .001) and posterior features (76% vs. 71%; P = .0046), While the difference of calcification was not significantly different.

Discussion: In this study, we aimed to prospectively verify a novel AI tool based on ultrasound dynamic videos and ACR BI-RADS characteristics. The prospective assessment suggested that the AI tool not only meets the clinical needs better but also reaches the diagnostic efficiency of experienced radiologists.

KEYWORDS

artificial intelligence, diagnosis, ultrasound video, BI-RADS, breast

1 Introduction

According to the latest statistics on cancer incidence and mortality from the International Agency for Research on Cancer (GRIBOCAN) in 2020, breast cancer incidence has risen to the top and become the first cause of death among women worldwide (1). Early screening for breast cancer is crucial to reducing death rates. National guidelines for breast cancer screening vary from country to country. Due to the high proportion of dense breasts in Chinese women and the low sensitivity of mammography, the National Cancer Centre of China proposes that the ultrasound (US) should be the preferred method for breast cancer screening in Chinese women and recommends that women over 45 years old should be screened by ultrasound alone every 1-2 years (2). China has a large population which brings the heavy workload of breast cancer ultrasound screening. Therefore, it is necessary to develop a clinical application AI tool that can assist in diagnosis quickly and efficiently.

In clinical practice, to improve the accuracy of diagnosis, standardize ultrasound description, and communicate effectively with the physician, worldwide radiologists generally use the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) lexicon for breast US (3). The radiologist scans the whole breast with handheld ultrasound and gives BI-RADS category. However, since handheld ultrasound depends on the operators and experience, different radiologists have different opinions on the interpretation of BI-RADS characteristics, resulting in a high inter-observer variability, poor repeatability and low work efficiency (4–6).

To the best of our knowledge, AI is the most likely tool to improve diagnostic effectiveness and reduce the subjective and experience-dependent nature of handheld ultrasound. In recent years, with the continuous application of AI in clinics, deep learning has been favored by human experts due to its strong capacity for autonomous feature extraction and expression (7). Several studies applied deep learning to classify US images of breast lesions and have reported that it could achieve a high diagnostic performance similar to or better than that of experienced radiologists. Becker et al. (8) used a deep neural network to identify malignant lesions in 637 breast lesions. Han et al. (9) used the GoogLeNet convolutional neural network to classify benign and malignant ultrasound images of 7408 breast lesions. However, these studies are all based on a keyframe image which was not in accord with the actual situation of the clinical ultrasound dynamic scan. Moreover, a single static image cannot contain all the information about the entire breast lesion. In addition, studies (10, 11) showed that one person may also have other diagnoses within videos and static images for one lesion. Youk et al. (12) showed that in radiologists' interpretation of BI-RADS characteristics, videos had a higher diagnostic performance than static images. What's more, the above studies all belong to the benign and malignant dichotomy, which is of little clinical guiding significance compared with the multi-classification of BI-RADS. Ciritsis et al. (13) and Qian et al. (14) tried to use deep learning to conduct multi-classification studies of BI-RADS on breast lesions. However, they all merged BI-RADS 4a, 4b, and 4c into category 4, which was not in line with clinical practice, and at the same time, they still failed to overcome the limitations of using static images.

To overcome the above limitations, we first proposed an approach to scan the breast lesions and record their ultrasound dynamic videos per unified criteria, obtaining ACR BI-RADS morphological characteristics, and realizing the BI-RADS category. Compared to traditional methods based on single-frame static images, it not only captures comprehensive and complete breast lesion information, avoiding missing lesion features in static images but also better suits clinical diagnostic scenarios. In this approach, we introduce an AI diagnostic model (hereafter, Auto BI-RADS), which includes a YOLOV5 network with improved attention mechanism and morphological image processing algorithms. Based on effectively screening, localizing, and capturing tumor lesions in breast ultrasound dynamic videos, Auto BI-RADS can obtain BI-RADS morphological characteristics, achieve BI-RADS category and make a benign or malignant prediction. In this study, we prospectively verified its performance through a comparative test.

2 Materials and methods

2.1 Study sample

The institutional review board approved this study, and the requirement to obtain informed consent was waived (approval number: B-2022-182). In the development of Auto BI-RADS, we include retrospective data from the First Affiliated Hospital of Shantou University Medical College (Guangdong, China) with a total of 480 pathologically proven lesions. Figure 1 presents the inclusion and exclusion criteria. A total of 480 ultrasound dynamic videos equivalent to 18122 static images comprised the training and validation sets at 3:1 (mean age, 45 years; range, 18–82 years, May 2019 to June 2022). In the testing study, the dataset was screened with the same criteria and included two hospitals: internal test set (mean age 45 years, range,19-76 years, First Ultrasound Department, First Affiliated Hospital of Shantou University Medical College[Hospital 1], July 2022 to March 2023, n = 228); and external test set (Mean age: 50 years old; range,26-73 years old; Ultrasound Department, Shantou Chaonan Mingsheng Hospital [Hospital 2], July 2022 to March 2023, n = 64). A flowchart describing the research process is shown in Figure 1. Baseline clinical pathologic data, including age, sex, pathologic findings, and US diagnosis reports, were derived from the medical records. US video data were recorded by two experienced radiologists per the criteria below (SM.Q., with 7-8 years of experience, Hospital 1; JH.W., with 11-12 years of experience, Hospital 2).

2.2 US examination

US examinations were performed with linear array transducers of real-time US systems. All patients in hospital 1 were examined with the following US scanners: Canon Toshiba (Japan, Aplio I800, L9-18 MHz), SIUI (China, Apogee 6800, L8-12 MHz), and Siemens (Germany, Acuson S3000, L9-12MHz). Patients in hospital 2 were examined with Siemens (Germany, Acuson Sequoia, L4-10MHz).

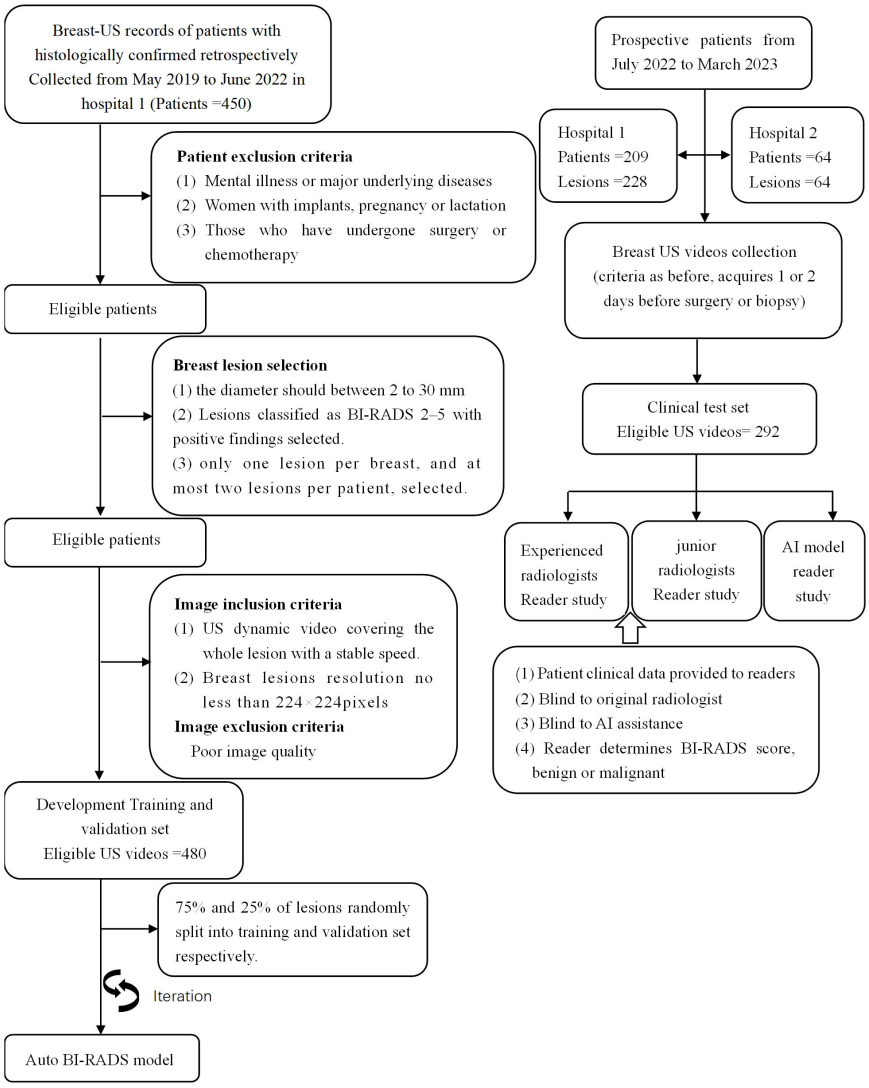


FIGURE 1
Overview of the retrospective and prospective workflow. It should be pointed out that all BI-RADS categories in this study were determined on 2D US videos exclusively.

US video acquisition: The patients held the supine position and raised their hands to fully expose their axilla and breast. We selected the maximum transverse diameter section of breast lesions and used the body mark showing the position; adjusted the depth to place the lesions in the center of the screen and the focus at the bottom of them; activated the storage function; kept the transducer at a constant speed to scan the lesions until some normal breast tissues appear, and pressed the storage key to acquire the video.

2.3 US image analysis

The six lexicon categories of BI-RADS were labeled as identifying features (shape, orientation, margin, echo pattern, posterior acoustic features, and calcification). The shape was a binary classification feature: regular or irregular; orientation was also binary: parallel or not parallel; the margin was another binary feature: circumscribed or not circumscribed. The echo pattern was mapped to three binary classification features: Anechoic, homogeneous echo (including homogeneous low, equal, and high echoes), heterogeneous echo (including heterogeneous solid and cystic-solid echoes); Posterior acoustic features are also classified into three classification features: no change, enhancement, or shadowing. Calcification is the last category, classified into three classification features: no calcification, coarse calcification, and punctate calcification (Figure 2).

In the training set, two experienced radiologists (DP.L. and XX.C, both with 10 years of US experience) were blinded to histopathologic results and independently manually labeled masks for each breast lesion video. Image classification was then performed based on the fifteen features of the six main BI-RADS lexicon. Groups would make a discussion to reach a consensus. In the validation dataset, the initial performance of Auto BI-RADS was evaluated by those two radiologists.

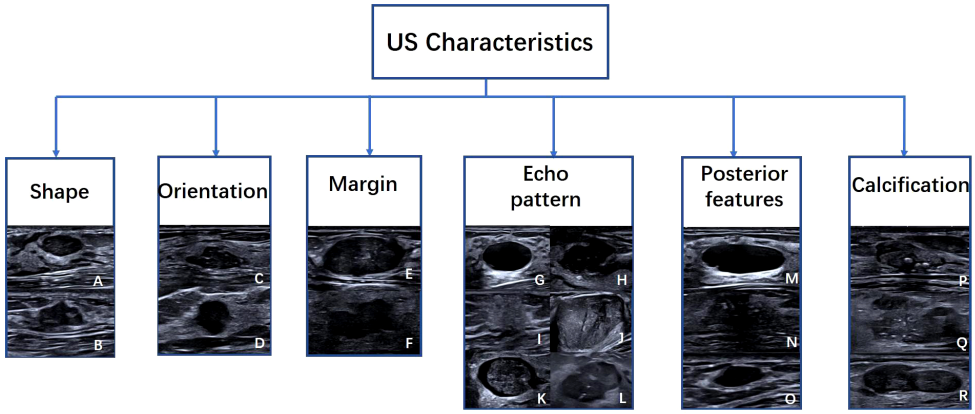


FIGURE 2
Examples of US images with Six American College of Radiology Breast Imaging Reporting and Data System characteristics. Characteristics include **(A)** regular, **(B)** irregular, **(C)** parallel, **(D)** not parallel, **(E)** circumscribed, **(F)** not circumscribed, **(G)** Anechoic, **(H)** hypoechoic, **(I)** hyperechoic, **(J)** Isoechoic, **(K)** complex cystic and solid, **(L)** heterogeneous, **(M)** enhancement, **(N)** shadowing, **(O)** no posterior features, **(P)** macrocalcifications, **(Q)** punctate echogenic foci, **(R)** no calcification.

For the test data set, four radiologists blinded to histopathologic results were split into two groups: experienced radiologists (BQ.Z. and XY.L., with 30 and 28 years of ultrasound work experience, respectively) and junior radiologists (ZY.L. and X.C., with 3 and 4 years of ultrasound work experience, respectively). Each radiologist independently evaluated the features of breast lesions in dynamic videos and determined the benign or malignant nature of the mass. When the evaluation results were inconsistent, a group consensus was reached through discussion.

2.4 AI model development

2.4.1 The establishment of Auto BI-RADS diagnosis model

The diagnostic AI model of Auto BI-RADS consists of three parts (Figure 3). The first part is based on the YOLOV 5 attention and segment network for object detection and segmentation. This model first converts the input breast lesions US video into sequence frames and selects frames containing lesions, then extracts and segments the regions of interest and their corresponding masks (15–18). In order to improve the detection performance of the network, we added the simple attention mechanism (Sim AM) to the model, which enhances the recognition effect of the small breast tumor target. We also combined the binary cross entropy loss (BCE loss) function, focal loss (FL loss) function, and Complete-Intersection-Over-Union (CIOU) loss function to optimize the network (19, 20). The second part focuses on extracting features of breast lesions using image processing algorithms. In this part, tumor regions of interest and their corresponding masks are obtained. These regions of interest and masks undergo equalization processing and data augmentation (21). After that, morphological image processing algorithms are used to extract features such as shape, orientation, margin, echo pattern, posterior acoustic features, and calcification from these tumor slices. The third part involves a feature score fusion algorithm based on weighted thresholds. Because multiple and unevenly distributed features may present in the breast lesions ultrasound video sequences, we establish threshold values based on the proportion of frames in which different features appear. Then, we merge the scores of all detected features and use a rank threshold score table to divide the tumor into its BI-RADS category and distinguish its benign or malignant nature.

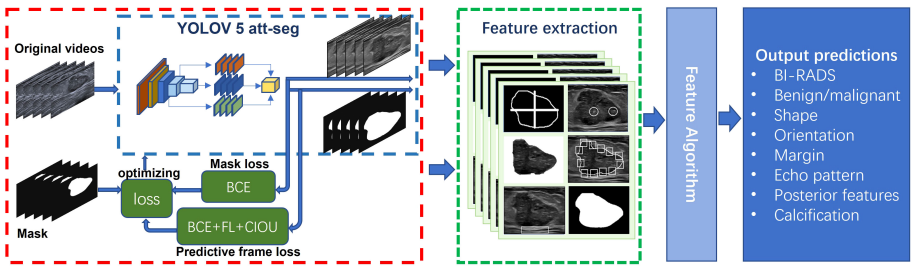


FIGURE 3
Proposed network scheme of the Auto BI-RADS model for breast lesions diagnosis based on American College of Radiology Breast Imaging Reporting and Data System US characteristics.

The predictive results of the Auto BI-RADS include the BI-RADS category, benign or malignant, and the assessment results of six significant features for each nodule. Figure 4 shows the prediction results of the Auto BI-RADS model in a case, compared with the interpretation results of experienced and junior radiologists.

2.4.2 The design of the YOLOV5 att-seg network based on the Sim AM

The deep learning model in this study is the YOLOV5 network based on the Sim AM. Hereafter we call it the YOLOV5 Attention Segmentation model (YOLOV5 att-seg).

The YOLOV5 network model consists of feature extraction and feature processing. The feature extraction part includes a Cross Stage Partial Network (CSP Net) (21) and a Path Aggregation Network (PA Net) (16) (Figure 5).

The CSP net is primarily composed of multiple Conv +BatchNorm+SiLU (CBS) modules, Cross Stage Partial modules 1 (CSP1), and spatial pyramid pooling fast (SPPF) modules. This network extensively utilizes residual structures and convolutional modules for refining image features and reducing feature map dimensions through downsampling. Additionally, it preserves feature maps at different depths within the network, allowing subsequent parts of the PA Net network to further integrate features from different levels.

The PA Net Network is primarily used for generating feature pyramids to enhance the model's detection of objects at different scales. It is an improvement based on the Feature Pyramid Network (FPN) architecture. The network consists of CBS modules, CSP2 modules, and Sim AM modules. Since CSP Net already captures sufficiently deep-level feature information, a non-residual module called CSP2 is used in the PA Net section to accelerate training and inference speed. The inclusion of Sim AM aims to further enhance the network's detection performance for small lesions. Sim AM is a parameter-free attention mechanism module based on the theory of neural energy functions (17). It calculates the neural energy of the input image and performs Hadamard multiplication with the input image to spontaneously enhance or suppress the neural pathways. (Figure 6) shows that YOLOV5 without Sim AM failed to identify tumor targets in some small breast tumor slices. However, the YOLOV5 att-seg model, which incorporates the Sim AM, exhibited improved detection performance for small tumor targets.

The second part is the feature processing section. It mainly involves processing the feature information obtained from the feature extraction section. The previous CSP Net and PA Net have effectively refined and aggregated the image features. Therefore, this part of the work is divided into two branches: mask segmentation and tumor target detection. One branch utilizes upsampling layers and CBS modules to further refine the edge information of the features, thereby obtaining better details of the mask edges. The other takes the three different-sized receptive field feature maps (RFFM) output by the PA Net network, adjusts them to the same dimension using CBS modules, and fuses them with the Concat function to enhance the network's detection of objects of different sizes. Finally, the network outputs the categories and detection boxes (22).

2.4.3 Performance validation of the YOLOV5 att-seg network model

The YOLOV5 att-seg network in this study was trained using 18122 static images extracted from 480 ultrasound dynamic videos. The detection box data was obtained by extracting the bounding rectangles from the masks. 75% of the samples were used for network training. The network was trained for 300 epochs, and the results are shown in (Figure 7).

In the remaining 25% samples for validation, we compared the detection performance of YOLOV5 att-seg with YOLOV5, Vgg-16 and Resnet50 networks. We also compared the segmentation performance of YOLOV5 att-seg with YOLOV5, Unet and Fcn-16s networks. The result is shown in Table 1.

The detection result shows that YOLOV5 att-seg has an improved performance in detecting smaller tumors in ultrasound images. In comparison, YOLOV5 att-seg vs. YOLOV5 vs. Vgg-16 vs. Resnet50, the precision, recall, and specificity are (0.98, 0.93, 0.94, vs. 0.97, 0.92, 0.9, vs. 0.84, 0.86, 0.82, vs. 0.78, 0.77, 0.77), respectively.

The segmentation result shows that YOLOV5 att-seg has an improved performance in precision, recall, Dice. and Iou. comparing with YOLOV5, Unet and Fcn-16s (0.98, 0.93, 0.77,0.68, vs. 0.98, 0.93, 0.75,0.67, vs. 0.60, 0.76, 0.62,0.53, vs. 0.53, 0.66, 0.53,0.46, respectively). However, the specificity of YOLOV5 att-seg (0.94) was slightly lower than that of Unet (0.97) and Fcn-16s (0.96). This indicates that the YOLOV5 att-seg model used in

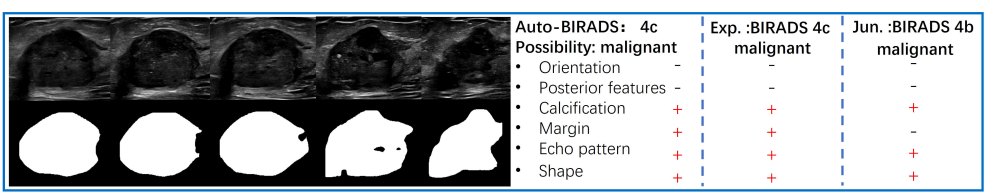


FIGURE 4
The left side displays a series of key frames extracted from an ultrasound dynamic video of a breast lesion that was pathologically proven as malignant. The right side shows the predictive results of the Auto BI-RADS model for the lesion based on ACR BI-RADS, as well as the interpretation results of the experienced (Exp.) and junior (Jun.) radiologists.

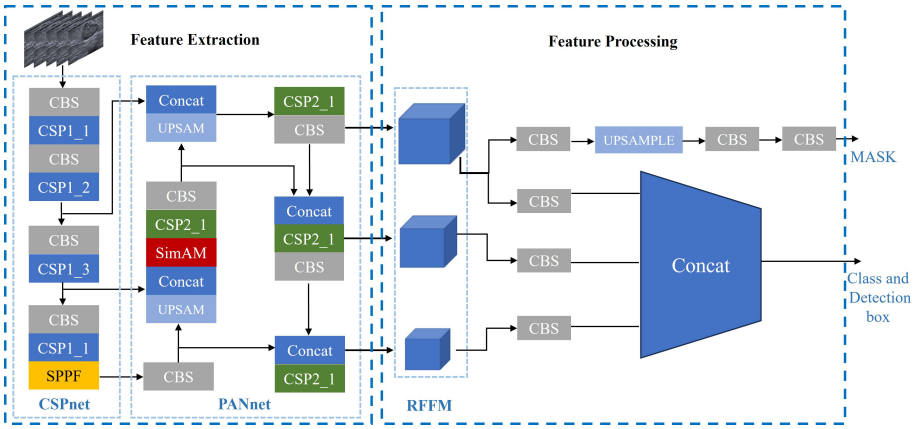


FIGURE 5
YOLOV5 att-seg Architecture.

this study achieves a more balanced performance compared to Unet and Fcn-16s. Unet and Fcn-16s tend to have overly conservative segmentation contours for ultrasound tumor targets, resulting in abnormally high specificity values. On the other hand, the enhancement in small target detection of YOLOV5 att-seg leads to improved segmentation performance compared to YOLOV5.

In terms of network running speed, YOLOV5 att-seg achieves significant improvement by using a single network to extract image features and obtain detection boxes and segmentation masks for tumor targets. This network demonstrates much faster speed compared to the traditional approach using separate networks for detection and segmentation.

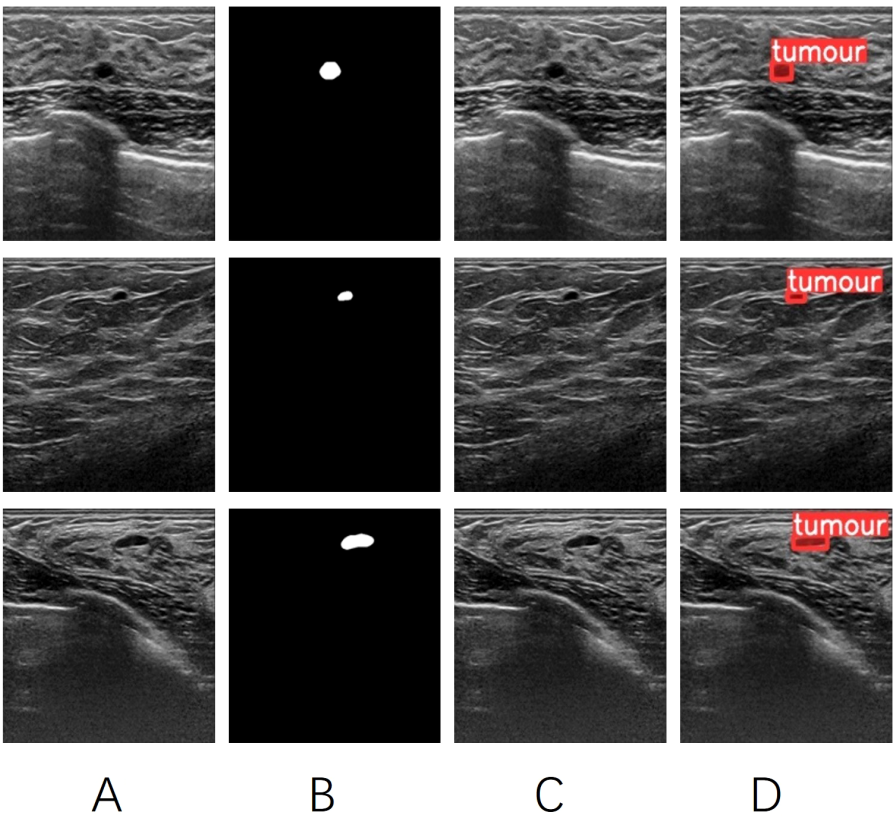


FIGURE 6
(A) the original image of a breast tumor. **(B)** the manually annotated ground truth by experienced physicians. **(C)** the detection results of YOLOV5 seg. **(D)** the detection results of YOLOV5 att-seg.

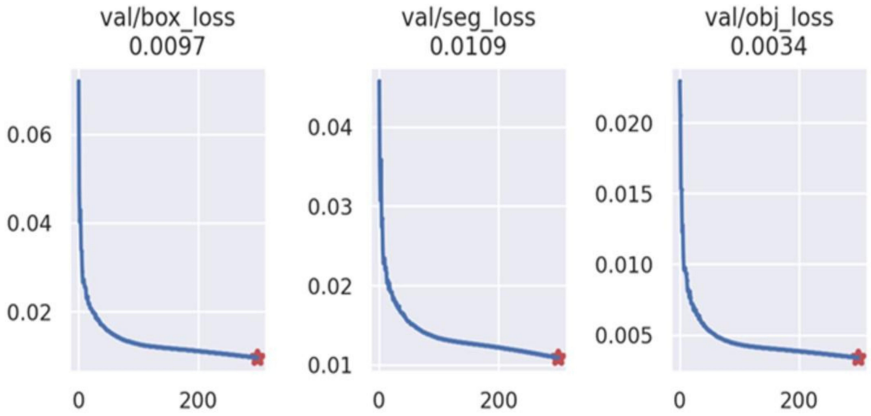


FIGURE 7
The box loss, segmentation loss, and target category loss of YOLOV5 att-seg achieve 0.0097, 0.0109, and 0.0034, respectively. It indicates that the network has achieved a good fit to the overall data.

2.5 Statistical analysis

The areas under the receiver operating characteristic curves (AUCs) with 95% confidence intervals (CI) were compared using the DeLong test (23) for Auto BI-RADS and two groups of radiologists. The threshold of Auto BI-RADS was established using validation sets. Performance metrics (sensitivity, specificity, positive predictive value, and negative predictive value) of Auto BI-RADS and the two groups of radiologists were evaluated. The Kappa test was used to compare the consistency of the breast lesion BI-RADS category between Auto BI-RADS and the two groups of radiologists. The McNemar test was used to compare the consistency rate of breast lesion characteristics recognition among Auto BI-RADS, the experienced group, and the junior group. Data were analyzed with SPSS, version 26.0 (IBM), and MedCalc, version 20.2 (MedCalc Software). $P<0.05$ was considered indicative of a statistically significant difference.

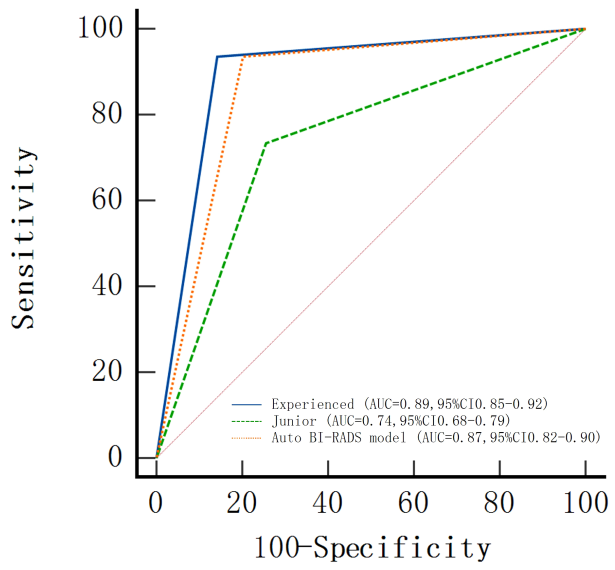


FIGURE 8
Areas under the receiver operating characteristic curves (AUCs) of the Auto BI-RADS model for breast lesions based on the American College of Radiology Breast Imaging Reporting and Data System and two radiologist groups with different experience levels who used BI-RADS.

3 Results

3.1 Patient characteristics and clinical features of breast lesions

A total of 698 patients were included in this study. In model development, 420 patients (480 pathologically confirmed lesions: 284 [60%] benign and 196 [40%)] malignant) from Hospital 1 were collected for training and validation in a 3:1 ratio. In the test data set, there were 278 patients (292 pathologically confirmed lesions: 168 [58%] benign and 124 [42%] malignant) from Hospitals 1 and 2. Figure 2 shows the workflow of patient inclusion and exclusion for model development and independent test. The specific pathological composition and distribution of the lesions are shown in Table 2.

3.2 Performance of Auto BI-RADS experienced radiologists and junior radiologists for diagnosing benign and malignant breast lesions

In the test set, the AUC value was slightly lower than that of the experienced group but significantly higher than that of the junior group as shown in Figure 8.

The AUC, sensitivity, specificity, positive predictive value, and negative predictive value of Auto BI-RADS were 0.87 (95%CI: 0.82, 0.90), 93% (116 out of 124 lesions), 81% (136 out of 168 lesions), 78% (116 out of 148 lesions), and 94% (136 out of 144 lesions), respectively. The false positive of the Auto BI-RADS is 6%, as much as the experienced group. We found no evidence of a statistical

TABLE 1 Comparison of performance metrics for object detection and segmentation in validation set.

Network	Function	Precision	Recall	Specificity	Dice.	Iou.
YOLOV5 att-seg	det./seg.	0.98|0.98	0.93|0.93	0.94|0.94	0.77	0.68
YOLOV5	det./seg.	0.97|0.98	0.92|0.93	0.91|0.91	0.75	0.67
Vgg-16	det.	0.84	0.86	0.82	–	–
Resnet50	det.	0.78	0.77	0.77	–	–
Unet	seg.	0.60	0.76	0.97	0.62	0.53
Fcn-16s	seg.	0.53	0.66	0.96	0.53	0.46

det, detection; seg, segmentation; Dice, Dice coefficient; Iou, Intersection over Union.

difference between the Auto BI-RADS model and the experienced group for AUC ($P = 0.06$), but there were statistically significant differences compared to the junior group ($P < 0.001$) (Table 3).

3.3 Comparison of consistency among Auto BI-RADS, experienced radiologists, and junior radiologists in the BI-RADS category in the test set

In the test set of 292 breast lesions, the consistency between the Auto BI-RADS model and the experienced radiologists in the BI-RADS category was higher than that of the junior radiologists, with kappa values of 0.82 and 0.60, respectively (Table 4).

3.4 Comparison of consistency rates among Auto BI-RADS, experienced radiologists, and junior radiologists in the identification of breast lesions characteristics in the test set

In the test set of 292 lesions, the consistency rate between Auto BI-RADS and experienced radiologists was higher than that between Auto BI-RADS and junior radiologists in the identification of morphology, orientation, margin, internal echo, posterior echo with respective values for morphology (93% [n = 271] vs. 80% [n = 234]; $P = 0.01$), orientation (90% [n = 265] vs. 84% [n = 247]; $P = 0.02$), margin (84% [n = 246] vs. 71% [n = 209]; $P = 0.01$), internal echo (69% [n = 202] vs. 56% [n = 163]; $P = 0.01$) and posterior echo (76% [n = 221] vs. 71% [n = 207]; $P = 0.046$). In the identification of calcification, there was no statistically significant difference in the consistency rates between Auto BI-RADS and experienced radiologists or junior radiologists ($P = 0.4$) (Table 5).

4 Discussion

In this study, we first developed a breast lesions AI classification model. By identifying the BI-RADS characteristics within the ultrasound dynamic videos, it can automatically evaluate the lesions' BI-RADS category and predict their benign or malignant nature.

The development of the Auto BI-RADS model was based on 480 breast lesions ultrasound videos equivalent to 18122 static images from Hospital 1, with a 3:1 ratio for training and validation. To verify the stability and efficiency of the model, we made an independent test in this study in 292 breast lesions testing sets from Hospital 1 and Hospital 2. Compared with those of experienced and junior radiologists, it showed that Auto BI-RADS achieved high performance in distinguishing between benign and malignant breast lesions (AUC: 0.87, sensitivity: 93%, specificity: 0.81), which was close to the experienced radiologists (AUC: 0.89, sensitivity: 93%, specificity: 86%), and significantly better than juniors (AUC: 0.74, sensitivity: 72%, specificity: 74%). Tracing back to the previous studies, Han et al. (9) first used an end-to-end deep learning framework to classify regions of interest selected by radiologists in a dataset of 7,408 static ultrasound breast lesions. They reported a sensitivity of 0.86, specificity of 0.93, and AUC >0.9. Ciritsis et al. (13) used a deep learning model that mimicked human decision-making to detect and classify ultrasound breast lesions in a dataset of 1,019 static images. In an external test dataset, they reported a sensitivity of 0.894, specificity of 1.0, and AUC of 0.967. Qian et al. (14) developed a neural network model that combined ultrasound B-mode and color Doppler to classify static ultrasound images of the breast in a larger dataset. Their bimodal model reported an AUC of 0.982, specificity of 88.7%, and sensitivity of 97%. Although the diagnostic performance indicators reported in those studies may appear higher than our study, they are not directly comparable. Firstly, the above studies were based on keyframes that can reflect the main BI-RADS characteristics of breast lesions. However, in clinical practice, not all radiologists were able to select the most critical frames. Secondly, a single static image cannot reflect all the morphological features of one breast lesion. Therefore, the results of those studies may have significant bias and low reproducibility, with limited clinical applicability. Additionally, none of the above studies conducted independent testing, raising questions about the stability of the models. In contrast, we used ultrasound dynamic videos for independent testing, which could improve its clinical generalizability with more objective and reproducible consequences.

In addition, we also compared the consistency among Auto BI-RADS, experienced radiologists, and junior radiologists in the

10.3389/fonc.2023.1274557

TABLE 2 Patient demographics data and breast lesion characteristics.

Characteristic	Retrospective Training and Validation sets	Prospective test sets	
		Hospital 1 Set	Hospital 2 Set
Number of patients	420	214	64
Age (years) (mean)	45 ± 12	45 ± 12	50 ± 11
Number of lesions	480	228	64
Lesion maximum diameter(mm)			
2-10	167(35%)	89(39%)	13(20%)
10-20	185(38%)	90(40%)	34(53%)
20-30	128(27%)	49(21%)	17(27%)
BI-RADS category [a]			
2	34(7%)	18(8%)	0(0%)
3	110(23%)	30(13%)	10(16%)
4a	106(23%)	70(31%)	8(13%)
4b	80(17%)	41(18%)	14(22%)
4c	74(15%)	35(15%)	22(34%)
5	76(16%)	34(15%)	10(15%)
Lesion type			
Invasive ductal carcinoma	82(17%)	26(11%)	14(22%)
Invasive lobular carcinoma	73(15%)	43(19%)	16(25%)
Ductal carcinoma *in situ*	21(4%)	10(4%)	6(9%)
Other malignant [b]	20(4%)	4(2%)	5(8%)
Fibroadenoma	141(30%)	64(28%)	13(20%)
Other benign [c]	143(30%)	81(36%)	10(16%)

The BI-RADS category [a] is based on the interpretation of the radiologist who originally performed the US examinations before the biopsy test, not the radiologists involved in the reader study. It should be noted that all BI-RADS categories involved in this study were determined on breast-US video images only. [b] Includes non-specific malignant results.
[c] Includes adenosis, hyperplasia, benign phyllodes tumors, and papillomata.

TABLE 3 Performance of Auto BI-RADS and two groups of radiologists for diagnosis of benign and malignant breast lesions in test set.

Parameter	Auto BI-RADS	Experienced Radiologists	Junior Radiologists
AUC	0.87[0.82,0.90]	0.89[0.85,0.92]	0.74 [0.68,0.79] †
Sensitivity (%)	93(116/124)	93(116/124)	72 (91/124)
Specificity (%)	81(136/168)	86(144/168)	74(125/168)
FP (%)	19(32/168)	14(24/168)	26(43/168)
FN (%)	6(8/124)	6(8/124)	25(33/124)
PPV (%)	78(116/148)	82(116/140)	68(91/134)
NPV (%)	94(136/144)	95(144/152)	79(125/158)

—Except where indicated, numbers in parentheses are numbers of lesions and 95% confidence intervals are in brackets.
FP, False Positive; FN, False Positive.
NPV, negative predictive value; PPV, positive predictive value.
† Data are for comparison with Auto BI-RADS (P < 0.001).

BI-RADS category (the Kappa values were 0.82 and 0.60, respectively). The results showed that Auto BI-RADS were highly consistent with experienced radiologists. Finally, we compared the identification of breast lesions' BI-RADS characteristics. The results showed that Auto BI-RADS had a higher consistency rate with experienced radiologists in morphology, orientation, margin, internal echo, and posterior echo. This indicates that the model conforms to the visual judgment of experienced human experts. These comparative studies have not been mentioned in previous studies. As for the recognition of calcification, there was no difference among the Auto BI-RADS model, experienced and junior radiologists. We speculate that this is because the characteristics of calcification are more complex, and their distribution in terms of location, size, and shape varies greatly. The image algorithm identifies different grayscale thresholds to determine whether calcification exists, and it will fail when there are only show slight changes in grayscale. It was also found that some tumors with high echogenic envelopes were mistakenly identified as calcifications. Chen et al. (24) had similar explanations in their identification of thyroid calcification. Later, we will increase the calcification samples and improve the algorithm to enhance the identification of calcification.

To meet the practical application, we first developed the Auto BI-RADS model based on ultrasound dynamic videos combining deep learning and image processing algorithms. There have been no similar reports previously. In the first step, we employed a YOLOV5 deep convolutional neural network to track and segment the targets. Then, we utilized image processing algorithms to extract BI-RADS features. Finally, we performed feature algorithm fusion to obtain target classification. For video tracking and segmentation, Yap et al. (25) have compared multiple types of deep learning neural networks, demonstrating their powerful capabilities in object tracking and segmentation. Relevant studies (26–28) have also indicated that deep learning exhibits uncertainty and a lack of interpretability in lesion feature recognition. Continuous learning required large sample sizes for the identification of each specific feature (29). However, machine learning has unique advantages in extracting breast lesion features. Hamyoo et al. (30) used machine learning alone to extract 13 features from lesions using the BI-RADS lexicon in a multi-center study (1288 static ultrasound images from three countries: Malaysia, Iran, and Turkey) and obtained an AUC value of 0.88, demonstrating the strong feature recognition capabilities of machine learning through comparison with human expert readings. Herein, our study constructed an Auto BI-RADS model based on deep learning and image processing algorithms to achieve the identification and classification of breast lesions in ultrasound dynamic videos. The prospective assessment indicates that the Auto BI-RADS model demonstrates good diagnostic performance and has significant potential.

Reviewing other imaging for breast cancer screening, mammography is considered a recommended method for reducing breast cancer-related mortality, but it involves radiation and is less

TABLE 4 Consistency rates between Auto BI-RADS model and two radiologist groups for classification of BI-RADS in test set.

US Characteristic	Rate between Auto BI-RADS and Experienced Radiologists (%)	Kappa Value	Rate between Auto BI-RADS and Junior Radiologists (%)	Kappa Value
BI-RADS 2	100(17/17)	0.82	82(14/17)	0.60
BI-RADS 3	97(36/37)		51(21/37)	
BI-RADS 4a	84(68/81)		77(63/81)	
BI-RADS 4b	80(42/52)		53(28/52)	
BI-RADS 4c	83(50/60)		70(42/60)	
BI-RADS 5	88(40/45)		75(34/45)	

-Numbers in parentheses are numbers of lesions (n = 292).
Cohen's Kappa coefficient is used to a diagnostic method. A Kappa value of 0.4–0.59 indicated weak agreement, 0.6–0.79 indicated moderate agreement, 0.8–0.9 indicated strong agreement and values above 0.9 indicated perfect agreement between two references.

sensitive to dense breasts, making it unsuitable for all countries (31). MRI is always used as a supplementary means (31). On the other hand, the US is recommended by Asia medical experts due to its low cost, non-radiation, and suitability for Asian women (31). The limitations of ultrasound have always been operator dependence and observer variability. Although many studies have focused on developing artificial intelligence models to address these limitations, they have not fully taken into account the practical clinical applications. Considering the above problems, we propose to develop a novel AI model simulating the clinical practice conducted by dynamic videos and BI-RADS characteristic identification. This approach allows for more objective, realistic, and reliable diagnostic results with high repeatability. The application of Auto BI-RADS offers great practical significance and provides better references for clinical practitioners with less experience.

Our study has several limitations. Firstly, the sample size for the prospective evaluation of the model is not large enough, and it does not include all categories of BIRADS, especially category 1 which indicates no lesion. Afterward, to improve the adaptability and stability of the model, we will include more external hospitals to increase the samples and species. By strengthening the model's training, we increase the model's robustness. Furthermore, it should be noted that the breast lesions ultrasound videos used in this study may still exhibit variability due to handheld ultrasound. In the future, if an Automated Breast Ultrasound System (ABUS) or robotic arms can be used to record the videos, it would provide more convincing results. Additionally, the latest fifth edition ACR BI-RADS guidelines have added color Doppler and elastography to evaluate breast lesions (32). It means that multimodal ultrasound has become part of breast cancer assessment. Therefore, for further improvement, this study can incorporate multiple ultrasound modalities such as color Doppler, elastography, contrast-enhanced ultrasound, etc., to develop a multimodal AI ultrasound diagnostic model.

TABLE 5 Consistency rates between Auto BI-RADS Model and two radiologist groups for identification of breast lesions characteristics in test set.

US Characteristic	Rate between Auto BI-RADS and Experienced Radiologists (%)	Rate between Auto BI-RADS and Junior Radiologists (%)	*P* Value
Shape	93(271)	80(234)	0.01
Orientation	90(265)	84(247)	0.02
Margin	84(246)	71(209)	0.01
Echo pattern	69(202)	56(163)	0.01
Posterior features	76(221)	71(207)	0.046
Calcification	31(91)	29(85)	0.4

-Numbers in parentheses are numbers of lesions (n = 292).

5 Conclusion

In conclusion, we first propose a novel method for breast tumor AI diagnosis based on breast lesions ultrasound dynamic videos to obtain ACR BI-RADS morphological characteristics, realize the BI-RADS category, and predict benign or malignant lesions. In the AI model development, we combined an improved attention mechanism YOLOV5 network with image processing algorithms to achieve it. This novel method not only avoids the problem of missing and incomplete lesion features caused by traditional single-frame static images but also better suits clinical diagnostic scenarios, providing a fast and effective approach for breast cancer screening.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of First Affiliated Hospital of Shantou University Medical College(approval number: B-2022-182). The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional

review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because the data are anonymous.

Author contributions

ZZ: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing – original draft, Writing – review & editing. SQ: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Resources, Writing – original draft. SZ: Methodology, Project administration, Software, Validation, Visualization, Writing – review & editing. BL: Investigation, Software, Validation, Writing – review & editing. JW: Data curation, Formal Analysis, Investigation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China under Grant (No. 82071992) and the Basic and Applied Basic Research Foundation of Guangdong Province (No. 2020B1515120061).

Acknowledgments

We thank Baoqun Zheng, MS, Xueying Li, BS, Xiaoxia, Chen, MS, Danpin Liu, MS, Xiao Chen, MS, Zeyin Lin, MS, Wenbin Rao, MS, Xinyu Lu, MS, Jingfeng Guo, MS, for their help throughout the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Ren W, Chen M, Qiao Y, Zhao F. Global guidelines for breast cancer screening: A systematic review. *Breast* (2022) 64:85–99. doi: 10.1016/j.breast.2022.04.003

3. Mendelson E, Böhm-Vélez M, Berg W, Whitman G, Feldman M, Madjar H. *Ultrasound in ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology (2013). p. 334.

4. Liu J, Zhou Y, Wu J, Li P, Liang X, Duan H, et al. Diagnostic performance of combined use of automated breast volume scanning & hand-held ultrasound for breast lesions. *Indian J Med Res* (2021) 154(2):347–54. doi: 10.4103/ijmr.IJMR_836_19

5. Brunetti N, Calabrese M, Martinoli C, Tagliafico AS. Artificial intelligence in breast ultrasound: from diagnosis to prognosis-A rapid review. *Diagn. (Basel)* (2022) 13 (1):1–17. doi: 10.3390/diagnostics13010058

6. Ibraheem SA, Mahmud R, Mohamad Saini S, Abu Hassan H, Keiteb AS, Dirie AM. Evaluation of diagnostic performance of automatic breast volume scanner compared to handheld ultrasound on different breast lesions: A systematic review. *Diagn. (Basel)* (2022) 12(2):541. doi: 10.3390/diagnostics12020541

7. Lee YJ, Choi SY, Kim KS, Yang PS. Variability in observer performance between faculty members and residents using breast imaging reporting and data system (BI-RADS)-ultrasound, fifth edition, (2013). *Iran J Radiol* (2016) 13(3):e28281. doi: 10.5812/iranjradiol.28281

8. Becker AS, Mueller M, Stoffel E, Marcon M, Ghafoor S, Boss A. Classification of breast cancer in ultrasound imaging using ageneric deep learning analysis software: a pilot study. *Br J Radiol* (2018) 91:20170576. doi: 10.1259/bjr.20170576

9. Han S, Kang HK, Jeong JY, Park MH, Kim W, Bang WC, et al. A deep learning framework for supporting the classification of breast lesions in ultrasound images. *Phys Med Biol* (2017) 62(19):7714–28. doi: 10.5812/iranjradiol.28281

10. Lee HJ, Kima EK, Kima MJ, Youk JH, Lee JY, Kang DR, et al. Observer variability of Breast Imaging Reporting and Data System (BI-RADS) for breast ultrasound. *Eur J Radiol* (2008) 65(2):293–8. doi: 10.1016/j.ejrad.2007.04.008

11. Abdullah N, Mesurolle B, El-Khoury M, Kao E. Breast Imaging Reporting and Data System lexicon for US: interobserver agreement for assessment of breast masses. *Radiology* (2009) 252:665–72. doi: 10.1148/radiol.2523080670

12. Youk JH, Jung I, Yoon JH, Kim SH, Kim YM, Lee EH, et al. Comparison of inter-observer variability and diagnostic performance of the fifth edition of BI-RADS for breast ultrasound of static versus video images. *Ultrasound Med Biol* (2016) 42 (9):2083–8. doi: 10.1088/1361-6560/aa82ec

13. Ciritsis A, Rossi C, Eberhard M, Marcon M, Becker AS, Boss A. Automatic classification of ultrasound breast lesions using a deep convolutional neural network mimicking human decision-making. *Eur Radiol* (2019) 29(10):5458–68. doi: 10.1007/s00330-019-06118-7

14. Qian X, Zhang B, Liu S, Wang Y, Chen X, Liu J, et al. A combined ultrasonic B-mode and color Doppler system for the classification of breast masses using neural network. *Eur Radiol* (2020) 30(5):3023–33. doi: 10.1007/s00330-019-06610-0

15. Redmon J, Divvala S, Girshick R, Farhadi A. (2016). You only look once: Unified, real-time object detection. *IEEE conference on computer vision and pattern recognition (CVPR)*, Las Vegas, NV, USA, 779–88. doi: 10.1109/CVPR.2016.91

16. Wang K, Liew JH, Zou Y, Zhou DQ, Feng JS. (2019). Panet: Few-shot image semantic segmentation with prototype alignment. *IEEE/CVF international conference on computer vision (ICCV)*, Seoul, Korea (South) 9197–205. doi: 10.1109/ICCV.2019.00929

17. Yang L, Zhang RY, Li L, Xie XH. (2021). Simam: A simple, parameter-free attention module for convolutional neural networks. *Proceedings of machine learning research (PMLR)* 139:11863–74.

18. Huang R, Ying Q, Lin Z, Zheng Z, Tan L, Tang G, et al. Extracting keyframes of breast ultrasound video using deep reinforcement learning. *Med Image Anal* (2022) 80:102490. doi: 10.1016/j.media.2022.102490

19. Lin TY, Goyal P, Girshick R, He K, Dollar P. (2020). Focal loss for dense object detection. *IEEE transactions on pattern analysis and machine intelligence* 42(2):318–27. doi: 10.1109/TPAMI.2018.2858826

20. Zheng ZH, Wang P, Liu W, Li J, Ye R, Ren D. (2020). Distance-IoU loss: Faster and better learning for bounding box regression. *Proceedings of the AAAI conference on artificial intelligence* 34(07):12993–3000. doi: 10.1609/aaai.v34i07.6999

21. Wang CY, Liao HYM, Wu YH, Chen PY, Hsieh JW, Yeh I-H. (2020). CSPNet: A new backbone that can enhance learning capability of CNN. *IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW)*. Seattle, WA, USA. 1571–80. doi: 10.1109/CVPRW50498.2020.00203

22. Bodla N, Singh B, Chellappa R. (2017). Soft-NMS–improving object detection with one line of code. *IEEE international conference on computer vision (ICCV)* 5562–70. doi: 10.1109/ICCV.2017.593

23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* (1988) 44(3):837–45. doi: 10.2307/2531595

24. Chen Y, Gao Z, He Y, Mai W, Li J, Zhou M, et al. An artificial intelligence model based on ACR TI-RADS characteristics for US diagnosis of thyroid nodules. *Radiology* (2022) 303(3):613–9. doi: 10.1148/radiol.211455

25. Yap MH, Pons G, Marti J, Ganau S, Sentis M, Zwiggelaar R, et al. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE J BioMed Health Inform* (2018) 22(4):1218–26. doi: 10.1109/JBHI.2017.2731873

26. McBee MP, Awan OA, Colucci AT, Ghobadi CW, Kadom N, Kansagra AP, et al. Deep learning in radiology. *Acad Radiol* (2018) 25(11):1472–80. doi: 10.1016/j.acra.2018.02.018

27. Stead WW. Clinical implications and challenges of artificial intelligence and deep learning. *JAMA* (2018) 320(11):1107–8. doi: 10.1001/jama.2018.11029

28. Chan HP, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis challenges and applications. *Adv Exp Med Biol* (2020) 1213:3–21. doi: 10.1007/978-3-030-33128-3

29. Tanaka H, Chiu SW, Watanabe T, Kaoku S, Yamaguchi T. Computer-aided diagnosis system for breast ultrasound images using deep learning. *Phys Med Biol* (2019) 64(23):235013. doi: 10.1088/1361-6560/ab5093

30. Hamyoon H, Yee Chan W, Mohammadi A, Yusuf Kuzan T, Mirza-Aghazadeh-Attari M, Leong WL, et al. Artificial intelligence, BI-RADS evaluation and morphometry: A novel combination to diagnose breast cancer using ultrasonography, results from multi-center cohorts. *Eur J Radiol* (2022) 157:110591. doi: 10.1016/j.ejrad.2022.110591

31. Mann RM, Hooley R, Barr RG, Moy L. Novel approaches to screening for breast cancer. *Radiology* (2020) 297(2):266–85. doi: 10.1148/radiol.2020200172

32. Spak DA, Plaxco JS, Santiago L, Dryden MJ, Dogan BE. BI-RADS((R)) fifth edition: A summary of changes. *Diagn Interv Imaging* (2017) 98(3):179–90. doi: 10.1038/s41551-021-00711-2

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 06 December 2023
DOI 10.3389/fonc.2023.1277556

OPEN ACCESS

EDITED BY
Junzhong Xu,
Vanderbilt University Medical Center, United States

REVIEWED BY
Yen-Peng Liao,
University of Texas Southwestern Medical Center, United States
Xiaoyu Jiang,
Vanderbilt University, United States

*CORRESPONDENCE
Sai Man Cheung
g.cheung@abdn.ac.uk

RECEIVED 14 August 2023
ACCEPTED 21 November 2023
PUBLISHED 06 December 2023

CITATION
Cheung SM, Wu W-S, Senn N, Sharma R, McGoldrick T, Gagliardi T, Husain E, Masannat Y and He J (2023) Towards detection of early response in neoadjuvant chemotherapy of breast cancer using Bayesian intravoxel incoherent motion. *Front. Oncol.* 13:1277556. doi: 10.3389/fonc.2023.1277556

Towards detection of early response in neoadjuvant chemotherapy of breast cancer using Bayesian intravoxel incoherent motion

Sai Man Cheung[1]*, Wing-Shan Wu[1], Nicholas Senn[1], Ravi Sharma[2], Trevor McGoldrick[2], Tanja Gagliardi[1,3], Ehab Husain[4], Yazan Masannat[5] and Jiabao He[1,6]

[1]Institute of Medical Sciences, School of Medicine, University of Aberdeen, Aberdeen, United Kingdom, [2]Department of Oncology, Aberdeen Royal Infirmary, Aberdeen, United Kingdom, [3]Department of Radiology, Royal Marsden Hospital, London, United Kingdom, [4]Department of Pathology, Aberdeen Royal Infirmary, Aberdeen, United Kingdom, [5]Breast Unit, Aberdeen Royal Infirmary, Aberdeen, United Kingdom, [6]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

Introduction: The early identification of good responders to neoadjuvant chemotherapy (NACT) holds a significant potential in the optimal treatment of breast cancer. A recent Bayesian approach has been postulated to improve the accuracy of the intravoxel incoherent motion (IVIM) model for clinical translation. This study examined the prediction and early sensitivity of Bayesian IVIM to NACT response.

Materials and methods: Seventeen female patients with breast cancer were scanned at baseline and 16 patients were scanned after Cycle 1. Tissue diffusion and perfusion from Bayesian IVIM were calculated at baseline with percentage change at Cycle 1 computed with reference to baseline. Cellular proliferative activity marker Ki-67 was obtained semi-quantitatively with percentage change at excision computed with reference to core biopsy.

Results: The perfusion fraction showed a significant difference (p = 0.042) in percentage change between responder groups at Cycle 1, with a decrease in good responders [–7.98% (–19.47–1.73), n = 7] and an increase in poor responders [10.04% (5.09–28.93), n = 9]. There was a significant correlation between percentage change in perfusion fraction and percentage change in Ki-67 (p = 0.042). Tissue diffusion and pseudodiffusion showed no significant difference in percentage change between groups at Cycle 1, nor was there a significant correlation against percentage change in Ki-67. Perfusion fraction, tissue diffusion, and pseudodiffusion showed no significant difference between groups at baseline, nor was there a significant correlation against Ki-67 from core biopsy.

Conclusion: The alteration in tumour perfusion fraction from the Bayesian IVIM model, in association with cellular proliferation, showed early sensitivity to good responders in NACT.

Clinical trial registration: https://clinicaltrials.gov/ct2/show/NCT03501394, identifier NCT03501394.

KEYWORDS

diffusion, cellularity, microcirculation, perfusion fraction, pathological response

1 Introduction

Neoadjuvant chemotherapy (NACT) is increasingly used in breast cancer, evolving from originally downstaging inoperable breast tumours to allow surgical excision (1) to facilitating potential breast and axillae conservation (2). However, NACT not only is costly at an estimated £6,000 per patient for a typical six-cycle regimen of 5-fluorouracil/epirubicin/cyclophosphamide (FEC 100) in the National Health Service (3) but also often leads to adverse side effects and subsequent severe physical and emotional distress (4, 5). Although NACT improves rates of pathological complete response (pCR) (6, 7) and disease-free survival (7, 8), poor responders to NACT might receive earlier and timely mastectomy or breast conservation (9). RECIST criterion, the current approach to estimate residual disease load based on tumour size (10) at the halfway point of NACT (11), has limited accuracy at a relatively late stage of treatment, demanding more precise radiological approaches.

The loss of tumour cellularity is the central histological marker of cellular damage in tumours responding to NACT (12). Diffusion-weighted imaging (DWI), although sensitive to cellularity (13, 14) with the potential of identifying responders after one cycle of NACT (15), is susceptible to biological noise and limited to large cohort studies (16), and is therefore inadequate for response-guided NACT (17). Apparent diffusion coefficient (ADC) from DWI (18) is effective in differentiation of tumour from healthy tissue and benign lesions (19, 20). An increase in ADC at the halfway point of 12 weeks of NACT-predicted pCR, however, may not reach clinical relevance with the receiver operating characteristics curve at an area under 0.6 (21). Diffusion tensor imaging yielded a significant increase in prime diffusion coefficient (λ_1) and ADC in good responders compared to poor responders at the completion of NACT, although baseline diffusion metrics did not predict good response (22). Diffusion kurtosis imaging approximates the deviation from the tensor model using kurtosis, with a lower mean kurtosis at baseline associated with pCR at four cycles of NACT in patients with breast cancer (23). We have shown that *q*-space imaging was more effective in the evaluation of cellularity in breast cancer; however, the method was not suitable for routine clinical application due to the demand on high field gradient and long scan duration (24). Intravoxel incoherent motion (IVIM), incorporating tissue diffusion and blood microcirculation as two independent components (18), showed improved diagnostic sensitivity in breast cancer (25). However, IVIM is prone to misfitting as a result of the high susceptibility in the main algorithm to biological noise (26). A Bayesian probability (BP) approach has been suggested to improve fitting accuracy and reduce variability in the estimation of tissue diffusion and blood microcirculation (27).

We therefore hypothesise that the Bayesian IVIM model may differentiate good from poor responders at baseline and after Cycle 1 of NACT with association from tumour proliferative activity, providing a non-invasive biomarker sensitive to prediction and early response to NACT.

2 Materials and methods

We hence conducted a prospective, longitudinal study of NACT in 17 female patients with breast cancer using the Bayesian IVIM model (Figure 1). The study was approved by the London Research Ethics Committee (Identifier: 17/LO/1777) and registered as a clinical trial [NCT03501394]. The planned study incorporated four MRI scans across the entire NACT, but was interrupted and closed prematurely due to the COVID-19 pandemic. Therefore, analysis was conducted on MRI scans acquired at baseline and after Cycle 1 only.

Clinical Procedure: Seventeen female patients (age 37–71 years), with grade II or III invasive breast carcinoma from core biopsy and planned for NACT were recruited into the study. Patients with a previous history of breast cancer or receiving hormonal treatment were not eligible. All patients received 5-fluorouracil 500 mg/m^2, epirubicin 100 mg/m^2, and cyclophosphamide 500 mg/m^2 (FEC) once every 21 days for the first three cycles, and docetaxel 100 mg/m^2 once every 21 days for the remaining three cycles (28, 29). Two patients with HER2-positive breast cancer additionally received pertuzumab and trastuzumab for a year (30, 31). MRI scans were performed at 5–10 days (median: 7) before the start of the treatment and 10–14 days (median: 12) after Cycle 1. MRI was acquired from 17 patients at baseline and 16 patients at Cycle 1 due to complications in one patient. Standard clinical histopathological examination was performed for each patient to determine histological grade, and immunostaining of Ki-67, a nuclear marker of cellular proliferation associated with worse survival

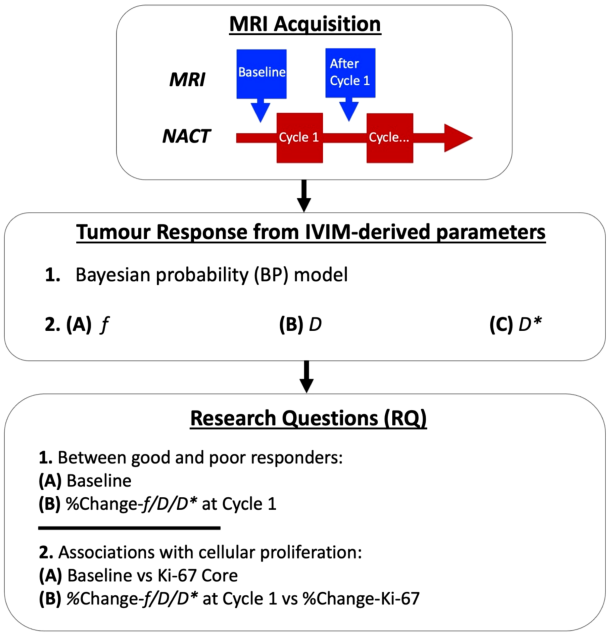


FIGURE 1
Intravoxel incoherent motion (IVIM) images were acquired before neoadjuvant chemotherapy (NACT) at baseline and after Cycle 1. A Bayesian probability (BP) IVIM model was used to compute perfusion fraction (*f*), tissue diffusion (*D*), and pseudodiffusion (*D**) for the assessment of prediction and early tumour response to NACT. The baseline and percentage change in *f*, *D*, and *D** at Cycle 1 were examined between good and poor responders, with patients grouped according to the Miller–Payne system for pathological response (RQ1). Medians of baseline and percentage change in *f*, *D*, and *D** were compared against tumour cellular proliferation marker Ki-67 at core biopsy and percentage change in Ki-67, respectively, from immunostaining in histopathology (RQ2).

outcomes (32), was conducted in a single batch. The histology results were obtained from core biopsies before NACT and resected residual tumours after six cycles respectively, with appropriate positive controls (33). The pathological response was assessed on the resected tumours, and the good responders and poor responders were identified as above (grades 4 and 5) or below (grades 1, 2, and 3) 90% reduction in cellularity, respectively, according to the Miller–Payne system (12). The percentage change in Ki-67 was computed as the difference between biopsy and excision, normalised to biopsy: [Ki-67 in resected tumour – Ki-67 in core biopsy]/Ki-67 in core biopsy × 100%.

Magnetic Resonance Imaging: All images were acquired on a 3 T clinical whole-body MRI scanner (Achieva TX, Philips Healthcare, Best, The Netherlands), using body coil for uniform transmission and a 16-channel breast coil for signal detection. Patients were in prone position with the imaging volume centered on the breast affected by tumour. IVIM images were acquired in the sagittal orientation using pulsed gradient spin echo (PGSE) sequence with single-shot echo planar imaging (EPI) at 10 diffusion weightings (*b*-values at 0, 30, 60, 90, 120, 250, 400, 600, 800, and 1,000 s/mm^2) (34). For each *b*-value, diffusion gradients were applied along three orthogonal directions, and the image was computed as the average across the three directions. Images were acquired with a diffusion time (δ/Δ) of 13.1/25.4 ms, a field of view (FOV) of 240 mm × 240 mm, an in-plane resolution of 2.5 mm × 2.5 mm, a slice thickness of 5 mm, an acceleration factor of 2, a repetition time (TR) of 2,400 ms, and an echo time (TE) of 50 ms.

Image Analysis: Bayesian IVIM was performed in MATLAB (R2020a, Mathworks, Natick, MA, USA). The tumour was delineated on dynamic contrast-enhanced MRI by a consultant radiologist in ImageJ (v1.58k, National Institute of Health, Bethesda, MD, USA), with adjustment of image resolution to match IVIM images and conservative definition of tumour boundary to avoid the necrotic, hemorrhagic, and cystic areas. The size of the tumours was evaluated based on the longest diameter from the high-resolution dynamic contrast-enhanced (DCE)-MRI (21, 35, 36). The Bayesian algorithm estimated the joint posterior distribution using the Rician noise likelihood function and uniform joint prior distribution, based on previous literature for Bayesian IVIM model fitting (37). The Bayesian fitting used a Markov chain Monte Carlo setup with Gibbs sampling and Metropolis-Hastings algorithm to derive a marginalised parameter distribution. The step-length parameters were updated every 2,000 iterations, with a total of 20,000 iterations. The conventional IVIM analysis algorithms, including nonlinear least squares full fitting, segmented-unconstrained, and segmented-constrained (38), were also deployed in supplementation to the study (Supplementary Data: Appendix A). The correction for the noise floor (39) was not undertaken since the data have a sufficiently high signal-to-noise ratio (SNR), and the same consistent approach was adopted for all the longitudinal data. The median perfusion fraction (*f*), tissue diffusion (*D*), and pseudodiffusion (*D**) within the tumour, representing volume fraction between capillary blood and tissue water, mean diffusivity of the tissue, and vascular blood flow motion, respectively, were calculated for baseline and Cycle 1. The percentage change in perfusion fraction, diffusion, and pseudodiffusion at Cycle 1 was computed with reference to baseline: [Cycle1(*f*/*D*/*D**) – Baseline(*f*/*D*/*D**)]/Baseline(*f*/*D*/*D**) × 100% (34).

Statistical Analysis: Statistical analysis was performed using the *R* software (v3.6.3, The *R* Foundation for Statistical Computing, Vienna, Austria). The normality of the distribution was assessed using the Shapiro–Wilk test. The measures at baseline and percentage change at Cycle 1 of perfusion fraction, diffusion, and pseudodiffusion were compared between good and poor responder groups using Wilcoxon rank sum test to determine the prediction and early sensitivity of the markers. The correlation of perfusion fraction, diffusion, and pseudodiffusion at baseline against Ki-67 from core biopsy for treatment-naïve prognosis was performed using Spearman's rank correlation test. The percentage change in perfusion fraction, diffusion, and pseudodiffusion against percentage change in Ki-67 for treatment-altered prognosis was also performed using Spearman's test. A p-value < 0.05 was considered statistically significant.

3 Results

The patient demographics is shown in Table 1. Among the 17 patients, there were 8 good responders and 9 poor responders at

10.3389/fonc.2023.1277556

TABLE 1 Tumour characteristics of patients.

Characteristic	All (*n* = 17)	Good Responders (*n* = 8)	Poor Responders (*n* = 9)	*p*-value
Age	51 (46–58)	50 (38–59)	52 (47–58)	NS
Tumour size at baseline (mm)	32 (26–38)	38 (34–43)	29 (20–37)	NS
Tumour size changes at Cycle 1 (%)	−7.3 (−16.7 to 0.0)	−16.7 (−27.1 to −4.4)	−3.9 (−8.3 to 0.0)	NS
Histology				
Invasive ductal carcinoma	16	7	9	
Mixed ductal/lobular carcinoma	1	1	0	
Grade				
Grade II	1	1	0	
Grade III	16	7	9	
Hormonal receptor status				
Oestrogen receptor positive (ER+)	7	3	4	
Human epidermal growth factor receptor 2 positive (HER2+)	2	2	0	
Triple negative (TN)	8	3	5	

Tumour histology and hormonal receptor status grouped by the Miller–Payne system (Poor Responders: 1, 2, and 3; Good Responders: 4 and 5). Median (interquartile range, IQR) of age, tumour size, and size changes are shown.
NS, not significant.

baseline, and due to complications, 1 patient did not complete an MR scan at Cycle 1. There was no significant difference in age and tumour size at baseline between good and poor responders. There was no significant difference in the change in tumour size between good and poor responders at Cycle 1 (Table 1).

There was no significant difference in Bayesian perfusion fraction (p = 0.481), tissue diffusion (p = 0.743), and pseudodiffusion (p = 0.673) at baseline between good and poor responders (Table 2; Supplementary Figure A1). There was also no significant difference in perfusion fraction, tissue diffusion, and

TABLE 2 Comparison of IVIM-derived parameters between responder groups before and after the first cycle of NACT and the association with Ki-67.

IVIM-derived parameters	Baseline-*f*/*D*/*D**			%Change-*f*/*D*/*D**			Ki-67 correlations (*ρ* score, *p*-value)	
	All (*n* = 17)	Good Responder (*n* = 8)	Poor Responder (*n* = 9)	All (*n* = 16)	Good Responder (*n* = 7)[a]	Poor Responder (*n* = 9)	Core[b]	% Change[c]
f [d]	10.81 (8.89–11.71)	9.95 (8.13–11.73)	11.01 (9.34–11.71)	5.54 (−10.34 to 15.37)	−7.98 (−19.47 to 1.73)*	10.04 (5.09–28.93)*	0.180, 0.480	0.590, 0.042*
D	0.95 (0.88–1.27)	0.95 (0.84–1.51)	0.96 (0.91–1.19)	16.06 (2.61–32.31)	28.87 (6.98–33.77)	15.54 (0.79–25.42)	−0.350, 0.174	0.380, 0.217
D*	6.20 (4.38–9.04)	4.63 (4.03–12.83)	6.23 (6.17–8.41)	−15.50 (−19.05 to −2.14)	−16.12 (−17.05 to −3.50)	−14.89 (−24.98 to −4.12)	0.190, 0.474	0.530, 0.075

[a]One patient did not complete MR scan due to complications.
[b]Spearman's rank correlation test – baseline-f/D/D* vs. Ki-67 Core.
[c]Spearman's rank correlation test – %Change-f/D/D* vs. %Change-Ki-67.
[d]Units at baseline – f: percentage (%), D and D*: $\times 10^{-3}$ mm^2/s.
The baseline and percentage change in perfusion fraction (f), tissue diffusion (D), and pseudodiffusion (D*) in good responders and poor responders from the Bayesian probability (BP) IVIM model. The Spearman's rank correlation coefficients (ρ) for baseline IVIM-derived parameters against Ki-67 in core biopsy and percentage changes in IVIM-derived parameters against percentage change in Ki-67 are also shown. Values are presented as median (IQR). Statistically significant differences ($p < 0.05$) are marked with an asterisk (*).

pseudodiffusion at baseline between good and poor responders from full fitting and segmented analyses (Supplementary Table A1, Supplementary Figure A2). There was no significant correlation in Bayesian perfusion fraction (p = 0.480), tissue diffusion (p = 0.174), and pseudodiffusion (p = 0.474) at baseline against Ki-67 from core biopsy (Table 2, Supplementary Figure A1). There was also no significant correlation in perfusion fraction, tissue diffusion, and pseudodiffusion at baseline against Ki-67 from core biopsy from full fitting and segmented analyses (Supplementary Table A1, Supplementary Figure A3).

There was a significant difference (p = 0.042) in percentage change in Bayesian perfusion fraction between good and poor responders at Cycle 1, with a decrease in good responders [−7.98% (−19.47–1.73), n = 7] against an increase in poor responders [10.04% (5.09–28.93), n = 9] (Figure 2A, Table 2). There was no significant difference in percentage change in perfusion fraction between good and poor responders at Cycle 1 from full fitting and segmented analyses (Figure 3; Supplementary Table A1). There was a significant correlation in percentage change in Bayesian perfusion fraction (p = 0.042, Figure 4A, Table 2) against percentage change in Ki-67. There was no significant correlation in percentage change in perfusion fraction against percentage change in Ki-67 from full fitting and segmented analyses (Figure 5; Supplementary Table A1). There was no significant difference in percentage change in Bayesian tissue diffusion (p = 0.606, Figure 2B, Table 2) and pseudodiffusion (p = 0.918, Figure 2C, Table 2) between good and poor responders at Cycle 1. There was no significant correlation in percentage change in Bayesian tissue diffusion (p = 0.217, Figure 4B, Table 2) and pseudodiffusion (p = 0.075, Figure 4C, Table 2) against percentage change in Ki-67. There was also no significant difference in percentage change in tissue diffusion and pseudodiffusion between good and poor responders (Figure 3, Supplementary Table A1), nor was there correlation against percentage change in Ki-67 from full fitting and segmented analyses (Figure 5; Supplementary Table A1).

The parametric maps from IVIM analysis from a typical good and poor responder at baseline and Cycle 1 are shown in Figure 6. The Ki-67-stained microscopy slides from a typical good and poor responder at core biopsy and excision are shown in Figure 7.

4 Discussion

In this study, we investigated predictive and early response markers for NACT in breast cancer using perfusion fraction, diffusion, and pseudodiffusion derived from BP IVIM. We found that perfusion fraction showed a significant alteration between baseline and Cycle 1 in good responders compared to poor responders, and the alteration is correlated with the change in proliferative activity accumulated through the whole course of NACT across the cohort. However, we did not observe significant differences in alterations in diffusion or pseudodiffusion at Cycle 1 between groups or their correlation against change in proliferative activities. We further did not observe significant differences in imaging markers at baseline between groups, or any significant correlation against proliferative activities at baseline.

The imaging markers of tissue diffusion, perfusion fraction, and pseudodiffusion at baseline did not predict NACT response, indicating the absence of evidence to use tissue diffusion and perfusion at baseline to guide NACT. The results were in agreement with imaging markers of diffusion tensor imaging and ADC at baseline that did not have predictive value for pCR after eight cycles of NACT (22). The results also agreed with a recent study showing that pretreatment tissue diffusion, perfusion fraction, and pseudodiffusion from the segmented constrained model were not predictors of response in patients undergoing a comparable regimen of NACT (36). Diffusion and perfusion metrics estimate cellularity and angiogenesis, respectively, and the lack of a difference between responder groups indicated that a tumour with high cell density and vascular abnormality at initial presentation might not determine the effectiveness of NACT, despite an initial poorer prognosis. Diffusion and perfusion metrics showed no correlation with Ki-67 prior to NACT, indicating no direct correlation between imaging markers of cellularity and angiogenesis with treatment-naïve prognosis, although tissue sampling error could not be excluded.

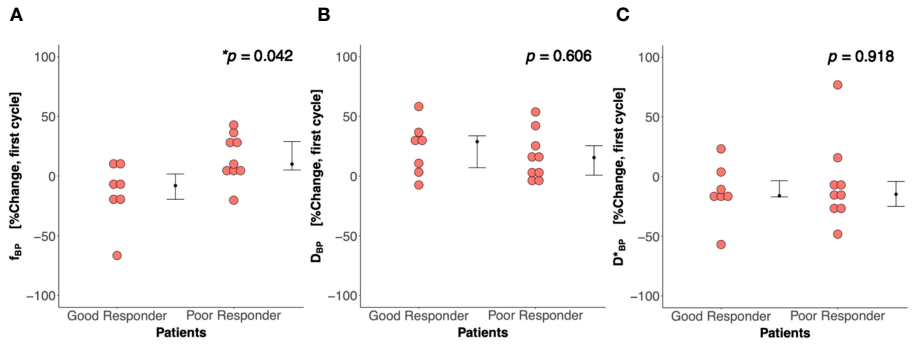


FIGURE 2
Percentage change in **(A)** perfusion fraction (*f*), **(B)** tissue diffusion (*D*), and **(C)** pseudodiffusion (*D**) between good and poor responders at first treatment cycle (Cycle 1) from the Bayesian probability (BP) IVIM model. There was a significant difference in percentage change in perfusion fraction between good and poor responders, but not in tissue diffusion and pseudodiffusion. Each dot represents the percentage change in *f*, *D*, and *D** from an individual patient. Error bar represents median (IQR). Statistically significant *p*-values (<0.05) are shown on the upper right corner with "*".

10.3389/fonc.2023.1277556

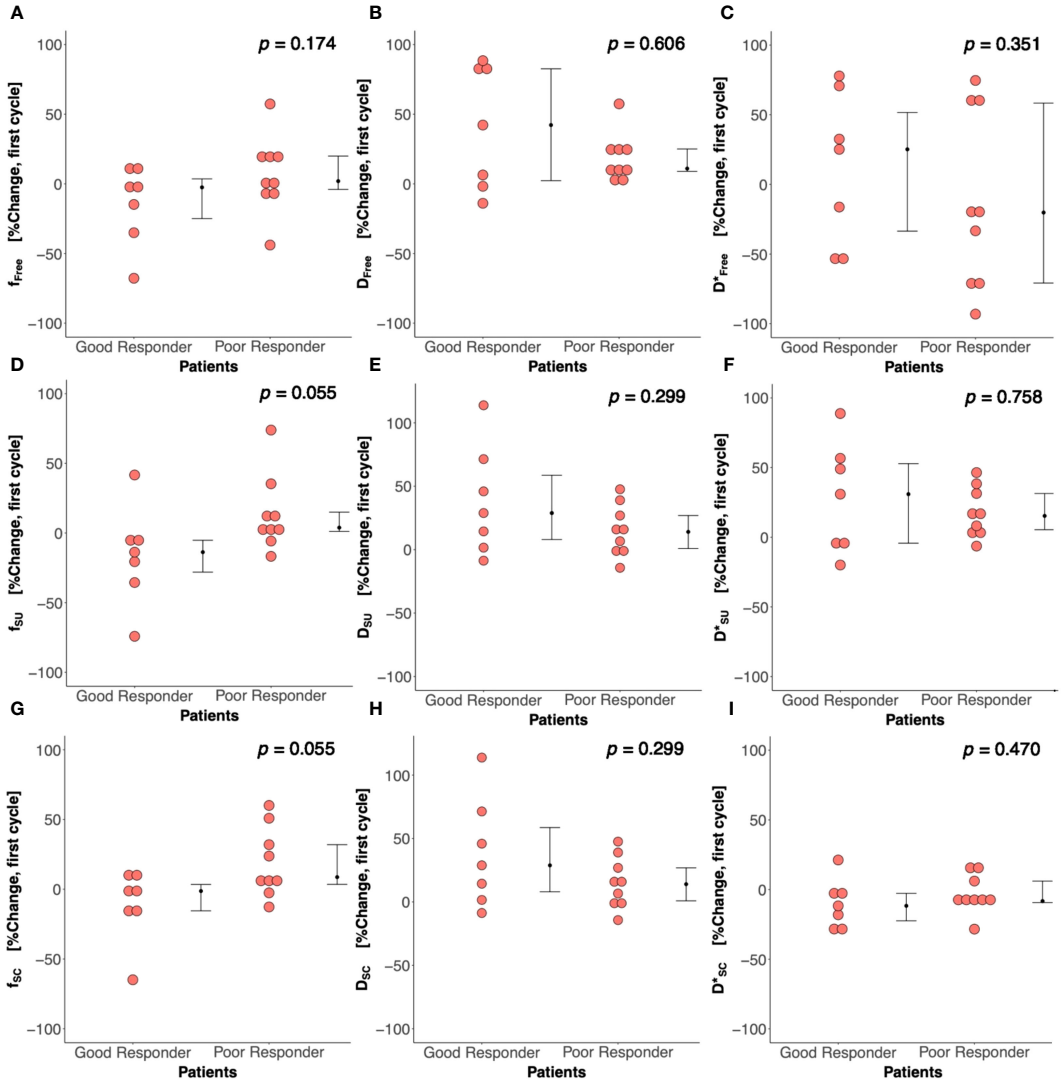


FIGURE 3
Percentage change in perfusion fraction (*f*), tissue diffusion (*D*), and pseudodiffusion (*D**) between good and poor responders from nonlinear least squares (Free), segmented-unconstrained (SU), and segmented-constrained (SC) IVIM models. The percentage change in *f*, *D*, and *D** between good and poor responders from **(A–C)** Free, **(D–F)** SU, and **(G–I)** SC algorithms are shown in dot plots. Each dot represents the IVIM-derived parameter of an individual patient. Error bar represents median (IQR).

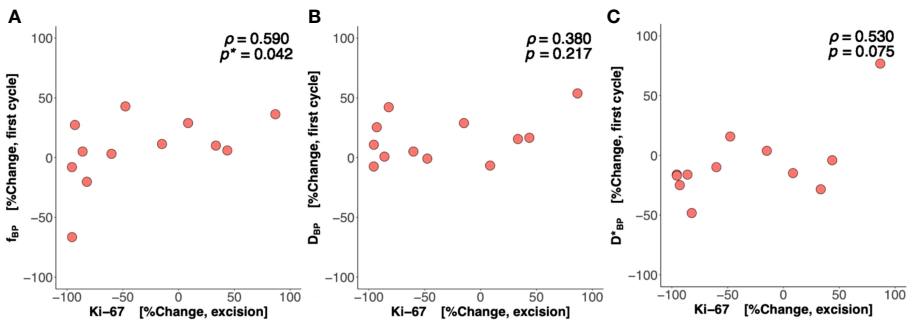


FIGURE 4
Percentage change in perfusion fraction (*f*), tissue diffusion (*D*), and pseudodiffusion (*D**) against percentage change in the tumour cellular proliferation marker Ki-67. The correlation of percentage change in **(A)** *f*, **(B)** *D*, and **(C)** *D** at Cycle 1 against percentage change in Ki-67 in resected tumour is shown in scatter plots. Spearman's rank correlation coefficient (*rho* (*ρ*)) was used for correlation analysis and respective *ρ* score and *p*-value are shown on each plot. Statistically significant *p*-values (<0.05) are marked by "*".

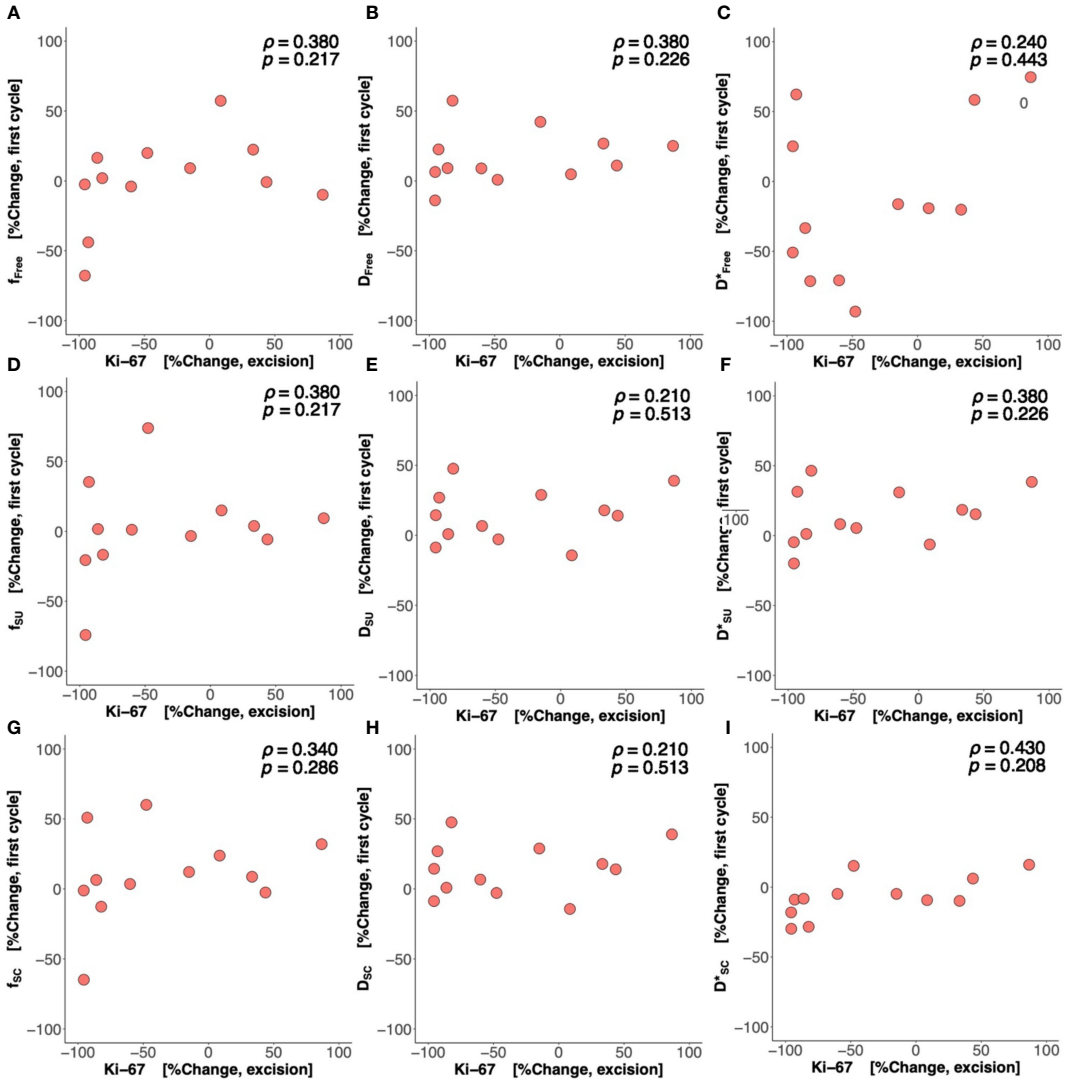


FIGURE 5
Percentage change in perfusion fraction (*f*), tissue diffusion (*D*), and pseudodiffusion (*D**) against percentage change in Ki-67. The correlations of percentage change in *f*, *D*, and *D** from **(A–C)** nonlinear least squares (Free), **(D–F)** segmented-unconstrained (SU), and **(G–I)** segmented-constrained (SC) algorithms at Cycle 1 against percentage change in tumour cellular proliferation marker Ki-67 in resected tumour are shown in scatter plots. Spearman's rank correlation coefficient (*rho* (*ρ*)) was used for correlation analysis and respective *ρ* score and *p*-value are shown on each plot.

There was an early significant decrease in perfusion fraction *f* in good responders, indicating that perfusion fraction might be a sensitive marker in the early identification of a successful NACT. The increase in stiffness of capillary vasculature obstructs microcirculation (40), leading to a modulation of perfusion in the tumour (41). Perfusion fraction has been shown to drastically decrease following a reduction in vascular blood flow motion, despite a subtle structural change in the functional capillary network (18). The susceptibility to systemic changes was lower in comparison to diffusion and pseudodiffusion as independent measures for the physiological response in cellularity and angiogenesis subsequent to cell apoptosis (42). Bayesian-derived perfusion fraction not only showed the potential of perfusion fraction as a marker to predict pathological complete response after one cycle of NACT, in agreement with a previous study (36), but also demonstrated a higher sensitivity since the full fitting and segmented analyses conducted in supplementation to Bayesian showed no group difference. The use of probability constraints on neighboring voxels in the Bayesian model led to less susceptibility of perfusion fraction and pseudodiffusion to the impact of noise, and improved the robustness of fitting (27). However, the higher demand on computing power may delay early adoption, whereas segmented analysis has an added advantage due to the faster processing time and initial validity in a recent study (36). Although perfusion volume ratio from DCE-MRI has been suggested as a marker of responders after one treatment cycle (43), DCE-MRI suffers from nonspecific contrast enhancement from post-treatment changes, including reactive inflammation, necrosis, and peritumoural oedema (18), requiring inputs from more than one radiologist (35). The sensitivity of DCE-MRI to

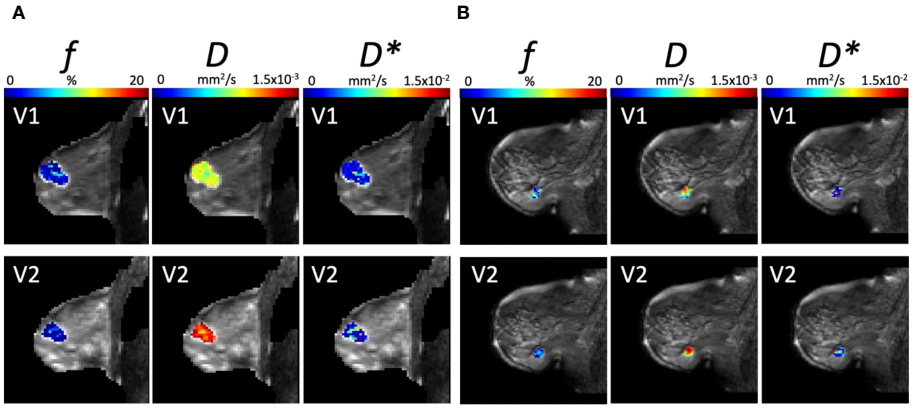


FIGURE 6
Parametric maps from IVIM Bayesian analysis of *f*, *D*, and *D** from a typical **(A)** good responder and **(B)** poor responder at baseline (V1) and Cycle 1 (V2) of NACT (overlaid on diffusion weighted images, b = 1000 s/mm^2). Images were acquired with a field of view of 240 mm × 240 mm, an in-plane resolution of 2.5 mm × 2.5 mm, a repetition time of 2,400 ms, and an echo time of 50 ms.

angiogenesis (44) is also limited by the accuracy in the measurement of arterial input function in kinetic hemodynamic models (45) and specialist quantitative deconvolution analysis (16). The inconsistency in terminology between radiology and research practice Dickie et al. 2023[1] further hinders the wider clinical adoption of quantitative perfusion maps from DCE-MRI for early response in NACT. IVIM, incorporating tissue diffusion and perfusion, shows clinical relevance in the current and previous studies (34, 36, 41, 46), does not require contrast, and has a clearer set of terminology to aid clinical translation. However, the higher susceptibility to noise demands extended acquisition time to reach submillimeter resolution sufficient for accurate determination of tumour size.

The significant correlation between percentage change in perfusion fraction after one cycle and percentage change in Ki-67 indicates a strong association between capillary blood-to-tumour water volume ratio with proliferative activities. Although a causal relationship for the primary impact of NACT on proliferative activity or blood supply could not be established, a reduction in metabolic demand from stunned proliferation and limitation of blood supply from restricted perfusion are both central characteristics of a successful NACT (46). The association between proliferative activity and perfusion has been shown in cell and *ex vivo* studies as central to tumour development (47, 48). Ki-67 was positively correlated with median (35) and mean (49) tumour perfusion fraction respectively in cross-sectional studies. Thus, an increase in proliferative activity has a corresponding increase in volume fraction between capillary blood and tissue water. Bayesian-derived perfusion fraction showed that good responders with a greater decrease in Ki-67 across NACT also had a greater decrease in perfusion fraction at one cycle, therefore enhancing the critical evidence in the clinical population from a longitudinal study. However, simultaneous full fitting and segmented analyses showed no correlation between change in perfusion fraction and Ki-67. A decrease from high pre-NACT (>35%) to low post-NACT (<15%) Ki-67 showed a sustained low recurrence (<20%) at 3 years after diagnosis (32), and post-NACT Ki-67 proliferative index is an independent prognostic marker in addition to pCR (32). The results showed the potential of perfusion fraction in early response for treatment-altered prognosis and the clinical relevance of an imaging biomarker in the targeted evaluation of the impact of NACT on breast tumours.

There was no significant difference in alteration of tissue diffusion *D* between responder groups, indicating that cellularity might not be the correct biological target to reveal the effectiveness of NACT at Cycle 1. It has been shown that an increase in tissue diffusion at the second (34, 41) and third (46) cycle was associated with good response in NACT; however, the time points are at a later stage of NACT and metabolic change at an earlier time point might precede morphological change in cellularity and hence tissue diffusion (34, 41, 46). There was a limited number of cytological or histological studies on changes in cellularity and metabolism during the early phase of NACT, potentially due to the heterogeneity across tumour and the fact that biopsy suffers from partial sampling error. There was a decrease in cellularity in biopsy obtained from good responders after two cycles of NACT (50), although the authors in the current study did not find any study on the direct assessment of cellularity after one cycle of conventional NACT. However, a low tumour cellularity in biopsy at day 15 in patients treated with anti-HER-based chemotherapy (including lapatinib and trastuzumab) (51) and a decrease in cellular proliferative activity of Ki-67 after one cycle of conventional NACT (52) predicted good responders. There was no significant difference in alteration of pseudodiffusion *D** between responder groups, in agreement with previous breast cancer treatment studies (34, 53). The results might be due to the higher variability in vascular blood flow motion within the capillary bed (54). The lack of association between alterations in tissue diffusion and

1 Dickie, B.R., Ahmed, Z., Arvidsson, J., Bell, L.C., Buckley, D.L., Debus, C., et al. (2023). A community-endorsed open-source lexicon for contrast agent-based perfusion MRI: A consensus guidelines report from the ISMRM Open Science Initiative for Perfusion Imaging (OSIPI). Magn Reson Med. Online ahead of print. doi: 10.1002/mrm.29840

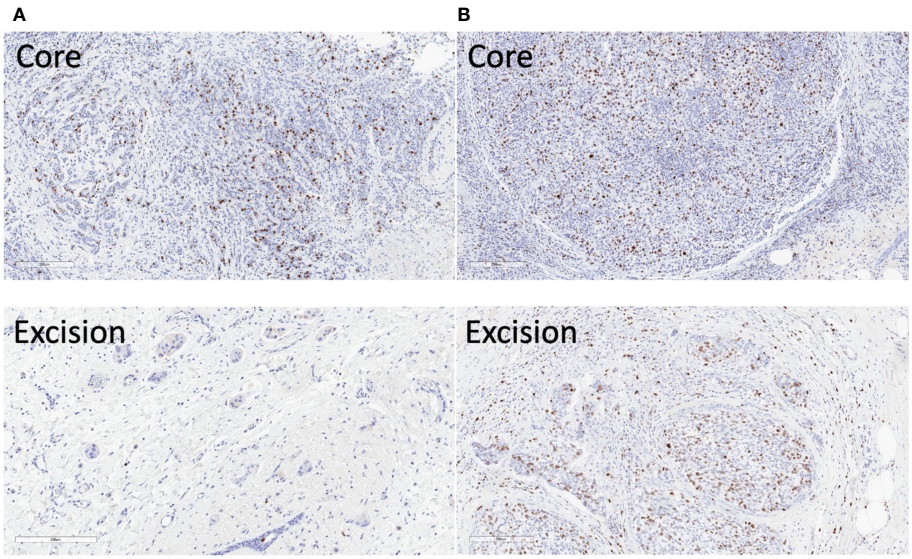


FIGURE 7
Ki-67 stained microscopy slides from a typical good and poor responder of neoadjuvant chemotherapy (NACT). **(A)** In the good responder, the Ki-67 score was 17.5% in the core biopsy and 0.8% in the resected tumour. **(B)** In the poor responder, the Ki-67 score was 23.7% in the core biopsy and 12.4% in the resected tumour. Sections at the greatest dimension of the specimens are shown. Magnification, ×10.

pseudodiffusion against alterations in Ki-67 showed an absence of evidence for a direct link between early response in cellularity and vascular blood flow motion against change in proliferative activity in the course of NACT.

Bayesian algorithm offers a robust assessment, and an improved estimation of perfusion fraction in association with pathology. The results of the study suggest that perfusion fraction might be a sensitive biomarker of NACT to improve treatment planning, reduce side effects, and expedite precision medicine. Mammography and breast ultrasound have been proposed at the halfway point of NACT to measure the residual tumour size using the RECIST criteria (55); however, tumour regression is not an accurate predictor of response at the first (56) or second (34) cycle of NACT. There was no correlation in size reduction with tumour grade decrease after two cycles of NACT (50), and a reduction in size of the tumours was seen in both small and large tumours (57), potentially due to the formation of islands of nonviable tumour cells subsequent to NACT (50). Perfusion fraction has the potential for tumour perfusion rate characterisation and responder identification after the first cycle, and the correlation with the change in Ki-67 showed that perfusion fraction might have a unique prognostic value in response-guided NACT prior to surgical intervention.

This investigation was a prospective, registered clinical trial that recruited consecutive patients, and set timing for individual MRI scans ensured comparability between patients (46). This study on patient data provided important clinical evidence to a previous study that used simulated and volunteer data (38) and showed that the Bayesian model might ensure greater accuracy of perfusion fraction in association with pathology for differentiation between good and poor responders. A threshold might not be clear cut, and hence, IVIM will contribute to NACT early responder identification but not as a standalone test. Future large cohort studies that will give an accurate estimation of sensitivity and specificity are required to demonstrate the potential of the Bayesian IVIM model to support early response markers in breast cancer management. A three-direction acquisition scheme was utilised due to limited acquisition time (58) and potential risk of overfitting with DTI parameters very sensitive to noise (39); however, a six-direction scheme (or more) might be used to mitigate the impact of anisotropy in the breast (39, 59, 60) in a future study. The current analysis might also benefit from a multi-compartmental IVIM model to account for the exchange between the extracellular and intracellular compartment that affects the quantification of diffusion and pseudodiffusion, since there was a characteristic change in cellular fibrous tissue (stroma) after NACT and the stromal component of the tumour is critical in tumour biology (50, 61).

5 Conclusion

The alteration in perfusion fraction from the Bayesian IVIM model supported the differentiation of good responders from poor responders at the first treatment cycle, and warrants further investigation in comparison to full fitting and segmented analyses in large cohort studies. Early treatment-induced changes in perfusion fraction might serve as non-invasive biomarker to facilitate the delivery of response-guided NACT and the development of an optimal treatment plan.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by London Research Ethics Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SMC: Data curation, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. W-SW: Formal Analysis, Investigation, Methodology, Writing – original draft. NS: Data curation, Investigation, Project administration, Writing – review & editing. RS: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing. TM: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing. TG: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing. EH: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Supervision, Writing – review & editing. YM: Conceptualization, Funding acquisition, Investigation, Supervision, Writing – review & editing. JH: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The study was jointly funded by the National Health Service Grampian Endowment Research Fund (16/11/047), Friends of Aberdeen and North Centre for Haematology, Oncology and Radiotherapy (RS17 004) and Tenovus Scotland (G16.09). SMC PhD study was jointly supported by Elphinstone scholarship, Roland Sutton Academic Trust and John Mallard scholarship and is currently funded by Cancer Research UK (C68628/A28312). NS PhD study was supported by Biotechnology and Biological Sciences Research Council (1654748, BB/M010996/1). The funding sources were not involved in the study design, in the collection, analysis and interpretation of data, in the writing of the report nor in the decision to submit the article for publication.

Acknowledgments

The authors would like to thank Dr. Matthew Clemence (Philips Healthcare Clinical Science, UK) for clinical scientist support, Ms. Erica Banks and Ms. Alison McKay for patient recruitment support, Ms. Teresa Morris and Ms. Dawn Younie for logistics support, Ms. Beverly McLennan, Ms. Nichola Crouch, Ms. Laura Reid, Mr. Mike Hendry for radiographer support, and Dr. Gordon Urquhart for providing access to the patients.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1277556/full#supplementary-material

References

1. Torrisi R, Marrazzo E, Agostinetto E, De Sanctis R, Losurdo A, Masci G, et al. Neoadjuvant chemotherapy in hormone receptor-positive/HER2-negative early breast cancer: When, why and what? *Crit Rev Oncol Hematol* (2021) 160:103280. doi: 10.1016/j.critrevonc.2021.103280

2. Banys-Paluchowski M, Gasparri ML, de Boniface J, Gentilini O, Stickeler E, Hartmann S, et al. Surgical management of the axilla in clinically node-positive breast cancer patients converting to clinical node negativity through neoadjuvant chemotherapy: current status, knowledge gaps, and rationale for the EUBREAST-03 AXSANA study. *Cancers (Basel)* (2021) 13(7):1565. doi: 10.3390/cancers13071565

3. NHS England and NHS Improvement. *National Cost Collection 2019/20 Report - National Schedule of NHS Costs*. London: NHS (2021).

4. Redana S, Sharp A, Lote H, Mohammed K, Papadimitraki E, Capelan M, et al. Rates of major complications during neoadjuvant and adjuvant chemotherapy for early breast cancer: An off study population. *Breast* (2016) 30:13–8. doi: 10.1016/j.breast.2016.07.019

5. Cain H, Macpherson IR, Beresford M, Pinder SE, Pong J, Dixon JM. Neoadjuvant therapy in early breast cancer: treatment considerations and common debates in practice. *Clin Oncol (R Coll Radiol)* (2017) 29(10):642–52. doi: 10.1016/j.clon.2017.06.003

6. von Minckwitz G, Untch M, Blohmer JU, Costa SD, Eidtmann H, Fasching PA, et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol* (2012) 30(15):1796–804. doi: 10.1200/JCO.2011.38.8595

7. Potter DA, Herrera-Ponzanelli CA, Hinojosa D, Castillo R, Hernandez-Cruz I, Arrieta VA, et al. Recent advances in neoadjuvant therapy for breast cancer. *Fac Rev* (2021) 10:2. doi: 10.12703/r/10-2

8. Kim MM, Allen P, Gonzalez-Angulo AM, Woodward WA, Meric-Bernstam F, Buzdar AU, et al. Pathologic complete response to neoadjuvant chemotherapy with trastuzumab predicts for improved survival in women with HER2-overexpressing breast cancer. *Ann Oncol* (2013) 24(8):1999–2004. doi: 10.1093/annonc/mdt131

9. Golshan M, Loibl S, Wong SM, Houber JB, O'Shaughnessy J, Rugo HS, et al. Breast conservation after neoadjuvant chemotherapy for triple-negative breast cancer: surgical results from the brighTNess randomized clinical trial. *JAMA Surg* (2020) 155 (3):e195410. doi: 10.1001/jamasurg.2019.5410

10. Bosch AM, Kessels AG, Beets GL, Rupa JD, Koster D, van Engelshoven JM, et al. Preoperative estimation of the pathological breast tumour size by physical examination, mammography and ultrasound: a prospective study on 105 invasive tumours. *Eur J Radiol* (2003) 48(3):285–92. doi: 10.1016/S0720-048X(03)00081-0

11. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026

12. Ogston KN, Miller ID, Payne S, Hutcheon AW, Sarkar TK, Smith I, et al. A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. *Breast* (2003) 12(5):320–7. doi: 10.1016/S0960-9776(03)00106-1

13. Baliyan V, Das CJ, Sharma R, Gupta AK. Diffusion weighted imaging: Technique and applications. *World J Radiol* (2016) 8(9):785–98. doi: 10.4329/wjr.v8.i9.785

14. Partridge SC, Nissan N, Rahbar H, Kitsch AE, Sigmund EE. Diffusion-weighted breast MRI: Clinical applications and emerging techniques. *J Magn Reson Imaging* (2017) 45(2):337–55. doi: 10.1002/jmri.25479

15. Gao W, Guo N, Dong T. Diffusion-weighted imaging in monitoring the pathological response to neoadjuvant chemotherapy in patients with breast cancer: a meta-analysis. *World J Surg Oncol* (2018) 16(1):145–y. doi: 10.1186/s12957-018-1438-y

16. Prevos R, Smidt ML, Tjan-Heijnen VC, van Goethem M, Beets-Tan RG, Wildberger JE, et al. Pre-treatment differences and early response monitoring of neoadjuvant chemotherapy in breast cancer patients using magnetic resonance imaging: a systematic review. *Eur Radiol* (2012) 22(12):2607–16. doi: 10.1007/s00330-012-2653-5

17. von Minckwitz G, Blohmer JU, Costa SD, Denkert C, Eidtmann H, Eiermann W, et al. Response-guided neoadjuvant chemotherapy for breast cancer. *JCO* (2013) 31 (29):3623–30. doi: 10.1200/JCO.2012.45.0940

18. Mendez AM, Fang LK, Meriwether CH, Batasin SJ, Loubrie S, Rodríguez-Soto AE, et al. Diffusion breast MRI: current standard and emerging techniques. *Front Oncol* (2022) 12:844790. doi: 10.3389/fonc.2022.844790

19. Chen X, Li WL, Zhang YL, Wu Q, Guo YM, Bai ZL. Meta-analysis of quantitative diffusion-weighted MR imaging in the differential diagnosis of breast lesions. *BMC Cancer* (2010) 10:693. doi: 10.1186/1471-2407-10-693

20. Chen L, Liu M, Bao J, Xia Y, Zhang J, Zhang L, et al. The correlation between apparent diffusion coefficient and tumor cellularity in patients: a meta-analysis. *PloS One* (2013) 8(11):e79008. doi: 10.1371/journal.pone.0079008

21. Partridge SC, Zhang Z, Newitt DC, Gibbs JE, Chenevert TL, Rosen MA, et al. Diffusion-weighted MRI findings predict pathologic response in neoadjuvant treatment of breast cancer: the ACRIN 6698 multicenter trial. *Radiology* (2018) 289(3):618–27. doi: 10.1148/radiol.2018180273

22. Furman-Haran E, Nissan N, Ricart-Selma V, Martinez-Rubio C, Degani H, Camps-Herrero J. Quantitative evaluation of breast cancer response to neoadjuvant chemotherapy by diffusion tensor imaging: Initial results. *J Magn Reson Imaging* (2018) 47(4):1080–90. doi: 10.1002/jmri.25855

23. Zhang D, Geng X, Suo S, Zhuang Z, Gu Y, Hua J. The predictive value of DKI in breast cancer: Does tumour subtype affect pathological response evaluations? *Magn Reson Imaging* (2022) 85:28–34. doi: 10.1016/j.mri.2021.10.013

24. Senn N, Masannat Y, Husain E, Siow B, Heys SD, He J. q-Space Imaging Yields a Higher Effect Gradient to Assess Cellularity than Conventional Diffusion-weighted Imaging Methods at 3.0 T: A Pilot Study with Freshly Excised Whole-Breast Tumors. *Radiol Imaging Cancer* (2019) 1(1):e190008. doi: 10.1148/rycan.2019190008

25. Liu C, Liang C, Liu Z, Zhang S, Huang B. Intravoxel incoherent motion (IVIM) in evaluation of breast lesions: comparison with conventional DWI. *Eur J Radiol* (2013) 82(12):782. doi: 10.1016/j.ejrad.2013.08.006

26. Orton MR, Collins DJ, Koh DM, Leach MO. Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven Bayesian modeling. *Magn Reson Med* (2014) 71(1):411–20. doi: 10.1002/mrm.24649

27. Vidić I, Jerome NP, Bathen TF, Goa PE, While PT. Accuracy of breast cancer lesion classification using intravoxel incoherent motion diffusion-weighted imaging is improved by the inclusion of global or local prior knowledge with bayesian methods. *J Magn Reson Imaging* (2019) 50(5):1478–88. doi: 10.1002/jmri.26772

28. Roché H, Fumoleau P, Spielmann M, Canon JL, Delozier T, Serin D, et al. Sequential adjuvant epirubicin-based and docetaxel chemotherapy for node-positive breast cancer patients: the FNCLCC PACS 01 Trial. *J Clin Oncol* (2006) 24(36):5664–71. doi: 10.1200/JCO.2006.07.3916

29. Zaheed M, Wilcken N, Willson ML, O'Connell DL, Goodwin A. Sequencing of anthracyclines and taxanes in neoadjuvant and adjuvant therapy for early breast cancer. *Cochrane Database Syst Rev* (2019) 2(2):CD012873. doi: 10.1002/14651858.CD012873.pub2

30. Gianni L, Pienkowski T, Im YH, Tseng LM, Liu MC, Lluch A, et al. 5-year analysis of neoadjuvant pertuzumab and trastuzumab in patients with locally advanced, inflammatory, or early-stage HER2-positive breast cancer (NeoSphere): a multicentre, open-label, phase 2 randomised trial. *Lancet Oncol* (2016) 17(6):791–800. doi: 10.1016/S1470-2045(16)00163-7

31. Liedtke C, Thill M, Jackisch C, Thomssen C, Müller V, Janni W, et al. AGO recommendations for the diagnosis and treatment of patients with early breast cancer: update 2017. *Breast Care (Basel)* (2017) 12(3):172–83. doi: 10.1159/000477575

32. von Minckwitz G, Schmitt WD, Loibl S, Müller BM, Blohmer JU, Sinn BV, et al. Ki67 measured after neoadjuvant chemotherapy for primary breast cancer. *Clin Cancer Res* (2013) 19(16):4521–31. doi: 10.1158/1078-0432.CCR-12-3628

33. Cheung SM, Husain E, Mallikourti V, Masannat Y, Heys S, He J. Intra-tumoural lipid composition and lymphovascular invasion in breast cancer via non-invasive magnetic resonance spectroscopy. *Eur Radiol* (2021) 31(6):3703–11. doi: 10.1007/s00330-020-07502-4

34. Kim Y, Kim SH, Lee HW, Song BJ, Kang BJ, Lee A, et al. Intravoxel incoherent motion diffusion-weighted MRI for predicting response to neoadjuvant chemotherapy in breast cancer. *Magn Reson Imaging* (2018) 48(5):27–33. doi: 10.1016/j.mri.2017.12.018

35. You C, Li J, Zhi W, Chen Y, Yang W, Gu Y, et al. The volumetric-tumour histogram-based analysis of intravoxel incoherent motion and non-Gaussian diffusion MRI: association with prognostic factors in HER2-positive breast cancer. *J Transl Med* (2019) 17(1):182–6. doi: 10.1186/s12967-019-1911-6

36. Almutlaq ZM, Wilson DJ, Bacon SE, Sharma N, Stephens S, Dondo T, et al. Evaluation of monoexponential, stretched-exponential and intravoxel incoherent motion MRI diffusion models in early response monitoring to neoadjuvant chemotherapy in patients with breast cancer-A preliminary study. *J Magn Reson Imaging* (2022) 56(4):1079–88. doi: 10.1002/jmri.28113

37. Jalnefjord O, Andersson M, Montelius M, Starck G, Elf AK, Johanson V, et al. Comparison of methods for estimation of the intravoxel incoherent motion (IVIM) diffusion coefficient (D) and perfusion fraction (f). *MAGMA* (2018) 31(6):715–23. doi: 10.1007/s10334-018-0697-5

38. Barbieri S, Donati OF, Froehlich JM, Thoeny HC. Impact of the calculation algorithm on biexponential fitting of diffusion-weighted MRI in upper abdominal organs. *Magn Reson Med* (2016) 75(5):2175–84. doi: 10.1002/mrm.25765

39. Iima M, Partridge SC, Le Bihan D. Six DWI questions you always wanted to know but were afraid to ask: clinical relevance for breast diffusion MRI. *Eur Radiol* (2020) 30(5):2561–70. doi: 10.1007/s00330-019-06648-0

40. Suresh S. Biomechanics and biophysics of cancer cells. *Acta Biomater* (2007) 3 (4):413–38. doi: 10.1016/j.actbio.2007.04.002

41. Che S, Zhao X, Ou Y, Li J, Wang M, Wu B, et al. Role of the intravoxel incoherent motion diffusion weighted imaging in the pre-treatment prediction and early response monitoring to neoadjuvant chemotherapy in locally advanced breast cancer. *Med (Baltimore)* (2016) 95(4):e2420. doi: 10.1097/MD.0000000000002420

42. Paran Y, Bendel P, Margalit R, Degani H. Water diffusion in the different microenvironments of breast cancer. *NMR Biomed* (2004) 17(4):170–80. doi: 10.1002/nbm.882

43. Tudorica A, Oh KY, Chui SY, Roy N, Troxell ML, Naik A, et al. Early prediction and evaluation of breast cancer response to neoadjuvant chemotherapy using quantitative DCE-MRI. *Transl Oncol* (2016) 9(1):8–17. doi: 10.1016/j.tranon.2015.11.016

44. Li L, Wang K, Sun X, Wang K, Sun Y, Zhang G, et al. Parameters of dynamic contrast-enhanced MRI as imaging markers for angiogenesis and proliferation in human breast cancer. *Med Sci Monit* (2015) 21:376–82. doi: 10.12659/MSM.892534

45. Petralia G, Summers PE, Agostini A, Ambrosini R, Cianci R, Cristel G, et al. Dynamic contrast-enhanced MRI in oncology: how we do it. *Radiol Med* (2020) 125 (12):1288–300. doi: 10.1007/s11547-020-01220-z

46. Bedair R, Priest AN, Patterson AJ, McLean MA, Graves MJ, Manavaki R, et al. Assessment of early treatment response to neoadjuvant chemotherapy in breast cancer using non-mono-exponential diffusion models: a feasibility study comparing the baseline and mid-treatment MRI examinations. *Eur Radiol* (2017) 27(7):2726–36. doi: 10.1007/s00330-016-4630-x

47. Baker E, Whiteoak N, Hall L, France J, Wilson D, Bhaskar P. Mammaglobin-A, VEGFR3, and ki67 in human breast cancer pathology and five year survival. *Breast Cancer (Auckl)* (2019) 13:1178223419858957. doi: 10.1177/1178223419858957

48. Ayoub NM, Jaradat SK, Al-Shami KM, Alkhalifa AE. Targeting angiogenesis in breast cancer: current evidence and future perspectives of novel anti-angiogenic approaches. *Front Pharmacol* (2022) 13:838133. doi: 10.3389/fphar.2022.838133

49. Feng W, Gao Y, Lu X, Xu Y, Guo Z, Lei J. Correlation between molecular prognostic factors and magnetic resonance imaging intravoxel incoherent motion histogram parameters in breast cancer. *Magn Reson Imaging* (2022) 85(1):262–70. doi: 10.1016/j.mri.2021.10.027

50. Sethi D, Sen R, Parshad S, Khetarpal S, Garg M, Sen J. Histopathologic changes following neoadjuvant chemotherapy in various Malignancies. *Int J Appl Basic Med Res* (2012) 2(2):111–6. doi: 10.4103/2229-516X.106353

51. Nuciforo P, Pascual T, Cortes J, Llombart-Cussac A, Fasani R, Pare L, et al. A predictive model of pathologic response based on tumor cellularity and tumor-infiltrating lymphocytes (CelTIL) in HER2-positive breast cancer treated with chemo-free dual HER2 blockade. *Ann Oncol* (2018) 29(1):170–7. doi: 10.1093/annonc/mdx647

52. Skoog L, Rutqvist LE, Wilking N. Analysis of hormone receptors and proliferation fraction in fine-needle aspirates from primary breast carcinomas during chemotherapy or tamoxifen treatment. *Acta Oncol* (1992) 31(2):139–41. doi: 10.3109/02841869209088893

53. Suo S, Yin Y, Geng X, Zhang D, Hua J, Cheng F, et al. Diffusion-weighted MRI for predicting pathologic response to neoadjuvant chemotherapy in breast cancer: evaluation with mono-, bi-, and stretched-exponential models. *J Transl Med* (2021) 19 (1):236–3. doi: 10.1186/s12967-021-02886-3

54. Lemke A, Laun FB, Simon D, Stieltjes B, SChad LR. An in *vivo* verification of the intravoxel incoherent motion effect in diffusion-weighted imaging of the abdomen. *Magn Reson Med* (2010) 64(6):1580–5. doi: 10.1002/mrm.22565

55. Keune JD, Jeffe DB, Schootman M, Hoffman A, Gillanders WE, Aft RL. Accuracy of ultrasonography and mammography in predicting pathologic response after neoadjuvant chemotherapy for breast cancer. *Am J Surg* (2010) 199(4):477–84. doi: 10.1016/j.amjsurg.2009.03.012

56. Pavlov MV, Bavrina AP, Plekhanov VI, Golubyatnikov GY, Orlova AG, Subochev PV, et al. Changes in the tumor oxygenation but not in the tumor volume and tumor vascularization reflect early response of breast cancer to neoadjuvant chemotherapy. *Breast Cancer Res* (2023) 25(1):12. doi: 10.1186/s13058-023-01607-6

57. Sheereen S, Lobo FD, Kumar B, Manoj Kumar S, Santosh Reddy G, Patel W, et al. Histopathological changes in breast cancers following neoadjuvant chemotherapy: implications for assessment of therapy-induced cytological and stromal changes for better clinical outcome and effective patient care. *Asian J Oncol* (2018) 04(02):61. doi: 10.1055/s-0038-1676909

58. Baltzer P, Mann RM, Iima M, Sigmund EE, Clauser P, Gilbert FJ, et al. Diffusion-weighted imaging of the breast-a consensus and mission statement from the EUSOBI International Breast Diffusion-Weighted Imaging working group. *Eur Radiol* (2020) 30 (3):1436–50. doi: 10.1007/s00330-019-06510-3

59. Partridge SC, Ziadloo A, Murthy R, White SW, Peacock S, Eby PR, et al. Diffusion tensor MRI: preliminary anisotropy measures and mapping of breast tumors. *J Magn Reson Imaging* (2010) 31(2):339–47. doi: 10.1002/jmri.22045

60. Teruel JR, Goa PE, Sjobakk TE, Ostlie A, Fjosne HE, Bathen TF. Diffusion weighted imaging for the differentiation of breast tumors: From apparent diffusion coefficient to high order diffusion tensor imaging. *J Magn Reson Imaging* (2016) 43 (5):1111–21. doi: 10.1002/jmri.25067

61. Honkoop AH, Pinedo HM, De Jong JS, Verheul HM, Linn SC, Hoekman K, et al. Effects of chemotherapy on pathologic and biologic characteristics of locally advanced breast cancer. *Am J Clin Pathol* (1997) 107(2):211–8. doi: 10.1093/ajcp/107.2.211

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 24 January 2024
DOI 10.3389/fonc.2024.1255618

OPEN ACCESS

EDITED BY
Jiabao He,
Newcastle University, United Kingdom

REVIEWED BY
Alessandro De Luca,
Sapienza University of Rome, Italy
Shuoyu Xu,
Bio-totem Pte Ltd, China
Jiang Zhu,
Shandong University, China

*CORRESPONDENCE
Yimin Zhang
dryiminzhang@163.com
Shengrong Sun
sun137@sina.com
Cheng Lei
leicheng@whu.edu.cn

†These authors have contributed equally to this work

RECEIVED 09 July 2023
ACCEPTED 08 January 2024
PUBLISHED 24 January 2024

CITATION
Zhang J, Deng J, Huang J, Mei L, Liao N, Yao F, Lei C, Sun S and Zhang Y (2024) Monitoring response to neoadjuvant therapy for breast cancer in all treatment phases using an ultrasound deep learning model. *Front. Oncol.* 14:1255618. doi: 10.3389/fonc.2024.1255618

Monitoring response to neoadjuvant therapy for breast cancer in all treatment phases using an ultrasound deep learning model

Jingwen Zhang[1†], Jingwen Deng[1†], Jin Huang[2†], Liye Mei[3], Ni Liao[1], Feng Yao[1], Cheng Lei[2,4,5*], Shengrong Sun[1*] and Yimin Zhang[1*]

[1]Department of Breast and Thyroid Surgery, Renmin Hospital of Wuhan University, Wuhan, China, [2]The Institute of Technological Sciences, Wuhan University, Wuhan, China, [3]School of Computer Science, Hubei University of Technology, Wuhan, China, [4]Suzhou Institute of Wuhan University, Suzhou, China, [5]Shenzhen Institute of Wuhan University, Shenzhen, China

Purpose: The aim of this study was to investigate the value of a deep learning model (DLM) based on breast tumor ultrasound image segmentation in predicting pathological response to neoadjuvant chemotherapy (NAC) in breast cancer.

Methods: The dataset contains a total of 1393 ultrasound images of 913 patients from Renmin Hospital of Wuhan University, of which 956 ultrasound images of 856 patients were used as the training set, and 437 ultrasound images of 57 patients underwent NAC were used as the test set. A U-Net-based end-to-end DLM was developed for automatically tumor segmentation and area calculation. The predictive abilities of the DLM, manual segmentation model (MSM), and two traditional ultrasound measurement methods (longest axis model [LAM] and dual-axis model [DAM]) for pathological complete response (pCR) were compared using changes in tumor size ratios to develop receiver operating characteristic curves.

Results: The average intersection over union value of the DLM was 0.856. The early-stage ultrasound-predicted area under curve (AUC) values of pCR were not significantly different from those of the intermediate and late stages ($p < 0.05$). The AUCs for MSM, DLM, LAM and DAM were 0.840, 0.756, 0.778 and 0.796, respectively. There was no significant difference in AUC values of the predictive ability of the four models.

Conclusion: Ultrasonography was predictive of pCR in the early stages of NAC. DLM have a similar predictive value to conventional ultrasound for pCR, with an add benefit in effectively improving workflow.

KEYWORDS

breast cancer, ultrasound image, deep learning, neoadjuvant chemotherapy, pathological complete response

Introduction

According to the 2023 cancer statistics from the American Cancer Society, breast cancer remains the most prevalent malignant tumor worldwide, and its incidence continues to rise (1). Thus, developing treatment and evaluation strategies remains crucial. Neoadjuvant chemotherapy (NAC) represents systemic medication administered before surgical tumor excision and is a standard treatment for locally advanced breast cancer (2). NAC can downstage tumors, rendering initially inoperable tumors eligible for surgery and enhancing the breast conservation rate (3). NAC can be used to evaluate tumor response to treatment by monitoring changes in tumor size during treatment (4).

Conventional assessment for tumor response including clinical examination, pathological examination and imaging examinations. The frequency of assessment of tumor response during NAC remains controversial. The National Comprehensive Cancer Network (NCCN) guidelines advocate for routine clinical examination to assess tumor response, with imaging evaluations only warranted if tumor progression is suspected (5). However, domestic guidelines recommend that imaging evaluations should be performed at least once every two cycles (6). Pathological examination post-NAC and surgery remains the gold standard for tumor response assessment (7). Pathological complete response (pCR) is the absence of residual invasive disease in the breast and axilla (8). Patients with pCR achieve long-term disease-free survival and improved overall survival rates (9, 10). Imaging examinations such as mammography, ultrasound, and magnetic resonance imaging (MRI) are employed to evaluate patients undergoing NAC (11, 12). Mammography can serve as an effective means for the primary tumor assessment and the detection of microcalcifications. Ultrasound provides real-time monitoring, is widely accessible and cost-effective (12, 13). Contrast-enhanced MRI is considered as the most sensitive imaging modality for assessing tumor response (14).

In addition to these conventional imaging examinations, artificial intelligence (AI) has been increasingly used to automatically improve early breast cancer detection and treatment. AI algorithms such as deep learning (DL) can efficiently and automatically analyze medical images, with outstanding capabilities in locating lesions and extracting characteristic features from medical images (15). DL has been introduced to assist clinicians in breast lesion identification and segmentation, cancer grading while allowing reproducibility and visualization (16–18). AI has been applied for ultrasound assessment of NAC treatment response in breast cancer (19). Ultrasound images exhibit speckle noise, variable tumor size and shape, and tumor-like breast tissue along with echo pattern modality imaging (20), which reduce diagnostic accuracy; therefore, developing precise tumor detection algorithms that lack noise and ambiguity represents considerable challenges.

However, certain aspects that require further exploration. First, the relationship between tumor size and NAC has long been a subject of research interest. Second, deep learning algorithms based on ultrasound images for predicting NAC response can be developed (21–23). Especially for breast ultrasound image segmentation, algorithms could be categorized into two methods. One algorithm is the CNN-based networks which utilize the fixed receptive region to extract information, such as U-net (24), FCN (25) and Mask R-CNN (26). Due to the special computer kernel, the networks pay more attention to the local features (27), which work poorly in evaluating the tumor-resemble, shadows and speckle noise. Though many researchers have creatively proposed many multi-scale and attention mechanisms, the improvement is limited. The other method utilized the transformer (28, 29), which splits the ultrasound image into tokens, which employ the sequence information to acquire the global relationship from the global dimension. Therefore, the transformer would extract more features especially avoiding the interference from tumor-resemble tissue. However, because the transformer most focuses on the global features, it detects boundaries of tumor would be sensitive. To enhance the accuracy of tumor detection, the network which balances the local and global features has challenges and has the potential to widespread in radiology (30). In comparison to the conventional manual delineation of tumor regions by sonographers, this model holds advantages in terms of segmentation speed and reduction of experiential bias. Moreover, existing studies primarily focus on early-stage treatment or predicting efficacy at individual time points (31–33). The continuous evaluation of treatment response of breast cancer during the process of NAC, rather than at isolated time points, remains to be investigated.

Herein, we prospectively collected data from patients with breast cancer who underwent NAC and designed a deep learning-based tumor detection model to analyze regions of interest (ROI) in ultrasound images. Next, we verified the ability of the model to detect breast tumors without noise interference and blurry boundaries using statistical methods to monitor ultrasound images during NAC, extract information from feature maps, and provide an intuitive and quantitative picture of tumor alterations. Furthermore, we examined the performance of proportional changes in tumor size measured using conventional models and DLM for predicting the response to chemotherapy. We employed DLM and conventional ultrasound throughout all phases of the NAC treatment to monitor its efficacy. The study aimed to assess the ability of DLM in predicting pCR in breast cancer patients undergoing NAC and its ability to evaluate the treatment response at various stages of NAC.

Methods

Patients

Between December 2020 and December 2022, researchers collected a total of 1393 ultrasound images from 913 patients at Renmin Hospital of Wuhan University. Among these, 956 ultrasound images from 856 patients were retrospectively collected and used as the training dataset. Additionally, 437 ultrasound images from 57 patients who underwent NAC were prospectively collected and utilized as the test dataset. This study was conducted according to the Declaration of Helsinki and relevant Chinese clinical trial research norms and regulations. Ethical approval was obtained from the Ethics Committee of

Renmin Hospital of Wuhan University (approval number: WDRY2022-K217). All patients provided written informed consent for ultrasound examinations, surgical intervention, and use of data. Figure 1 presents the patient selection flowchart.

Patient eligibility criteria included 1) a definitive diagnosis of primary invasive breast cancer by biopsy; 2) no previous treatment, and at least one of the following indications for NAC: tumor size > 5 cm; HER2 positive; estrogen receptor/progesterone receptor and human epidermal growth factor receptor (HER) 2 negative; axillary lymph node metastasis or strong breast-conserving intention; 3) underwent 6–8 cycles of complete NAC; 4) breast ultrasound examination before initiating NAC, after each NAC cycle, and after NAC completion; and 5) surgical resection after NAC completion.

Exclusion criteria were 1) patients who did not complete the NAC regimen or underwent treatment at another center; 2) breast surgery performed before NAC completion; 3) insufficient ultrasound image quality for feature extraction; and 4) lack of pathological results post-surgery.

Neoadjuvant chemotherapy regimen

The *National Comprehensive Cancer Network guidelines for Breast Cancer (Version 3.2020)* recommend selecting the NAC regimen based on the breast cancer molecular subtype. All patients underwent 6–8 NAC cycles. The most common regimen for patients with luminal or triple-negative breast cancer includes a combination of epirubicin, cyclophosphamide, and paclitaxel, administered every 21 days. Patients with HER2-positive tumors received either the THP (paclitaxel, trastuzumab, and pertuzumab every 21 days) or TCbHP (paclitaxel, carboplatin, trastuzumab, and pertuzumab every 21 days) regimens.

Ultrasound imaging

An ESOTE MEGAS GPX FD570A ultrasound diagnostic instrument was used for patient examination, with probe frequencies ranging 5–13 MHz. All patients were placed in the supine or side-lying position with both arms lifted and abducted to fully expose the breast. Diagnostic criteria were based on the American College of Radiology Breast Imaging Reporting and Data System. Experienced breast sonographers independently performed each ultrasonographic examination. Measurements were extracted from the ultrasound report and confirmed by another breast sonographer based on captured images. Two experienced breast sonographers manually segmented ROIs from original ultrasound images using a 3D slicer software (version 4).

Tumor response assessment

Tumor size was assessed using ultrasound after each NAC cycle. NAC-treated tumors were also evaluated using the Response

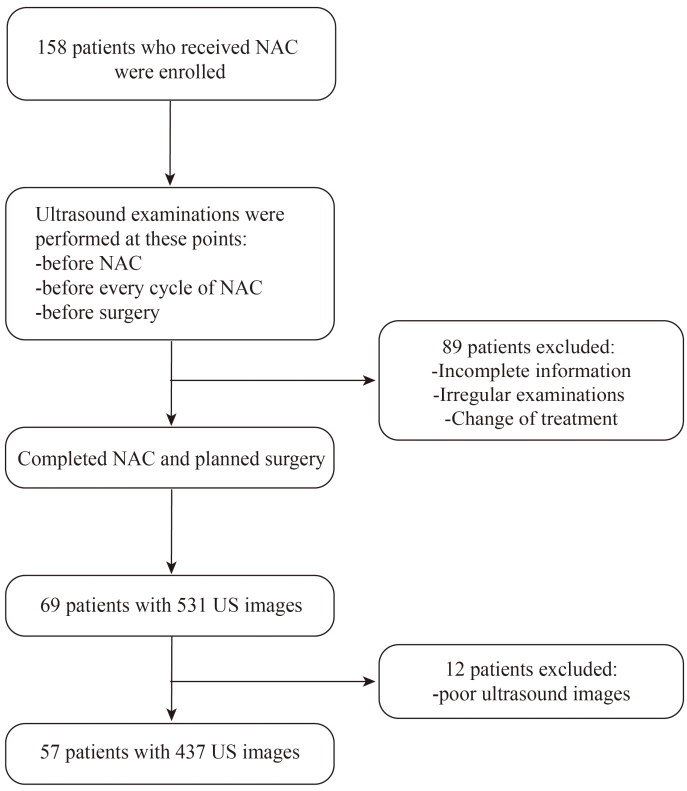


FIGURE 1
Flowchart of inclusion and exclusion for patients in the test set. US, ultrasound.

Evaluation Criteria for Solid Tumors (RECIST 1.1) (34). In this prospective study, tumor size was defined as the sum of the size of each tumor if the patient had multifocal disease. Four models were established to calculate tumor size before and after each treatment: the longest axis model (LAM), the product of two perpendicular axes model (dual-axis model [DAM]), the manual segmentation model (MSM), and the deep learning model (DLM). To compare changes in tumor size, the relative ratio was calculated after each treatment cycle using the following formula:

$$\text{Ratio N} = \frac{\text{untrasound tumor size after N cycles of NAC}}{\text{ultrsound tumor size before NAC}}$$

Histopathological assessment

Breast cancer was diagnosed by needle biopsy of the tumor. HER2 status was assessed using immunohistochemistry and fluorescence *in situ* hybridization analysis (35). Breast tumors were classified as HER2-negative and HER2-positive subtypes. After completing NAC, surgically resected breast tumor tissue was delivered to the Department of Pathology, where a specialized breast pathologist examined the specimens to establish the pathological diagnosis. The residual cancer burden (RCB) index was used as a criterion for assessing residual tumors after NAC for breast cancer. A pCR or RCB-0 was defined as the complete absence of invasive cancer in the breast and axillary lymph nodes (8).

Development of the deep learning-based model

Given that breast ultrasound images are characterized by low resolution and contrast and ambiguous boundaries, we utilized a custom U-Net neural network to capture tumor features by applying data augmentation (23), attention mechanism, and multi-scale method. As U-Net segments each pixel into classes, it can directly infer the ultrasound image and generate a tumor distribution map with the same dimensions as the input image. We adopted U-Net as a cancer-detection architecture and creatively added modules to construct a custom U-Net for enhancing extracted model features.

Figure 2 presents the algorithm construction process. To promote the generation and prevent overfitting, we adopted a data-augmentation technique and several network construction algorithms. During image processing, breast tumor features on the ultrasound image varied in all directions, lightness, and contrast; we utilized data-augmentation techniques such as cropping, rotation, and adjusting lighting conditions, including contrast and lightness, to extract general tumor features without intervention. We used batch normalization and dropout at each convolution layer in the neural network to avoid model oscillations and facilitate robustness. Additionally, we exploited the grayscale intensity as the input image, given that the grayscale ultrasound image contained sufficient information for the diagnosis, affording efficiency by reducing three channels into one. In the network, we introduced an attention mechanism after each convolution block, namely, the SCSE module (36), allocating more computing resources to abnormal regions and improving inference speed. Considering variations in breast tumor shape and size, a constant kernel size from the convolution layer was constrained to capture the tumor with a feasible receptive field. Therefore, we implemented multi-scale imaging to acquire different tumor sizes in ultrasound images, specifically atrous spatial pyramid pooling, with the capacity to extract multi-scale contextual information.

To ensure accurate tumor regions and contours, we proposed a hybrid loss function for model refinement. The loss function is crucial in selecting an optimizer for the model weight. We expanded the hybrid loss function for model adjustment based on the region and boundary. To obtain a more accurate intersection with the ground truth, we adopted the dice coefficient loss accompanying the binary cross-entropy loss to promote the accuracy of each pixel. Owing to the ambiguous tumor contour due to poor contrast, we adopted the active contour (37) and Hausdorff distance loss (38) to

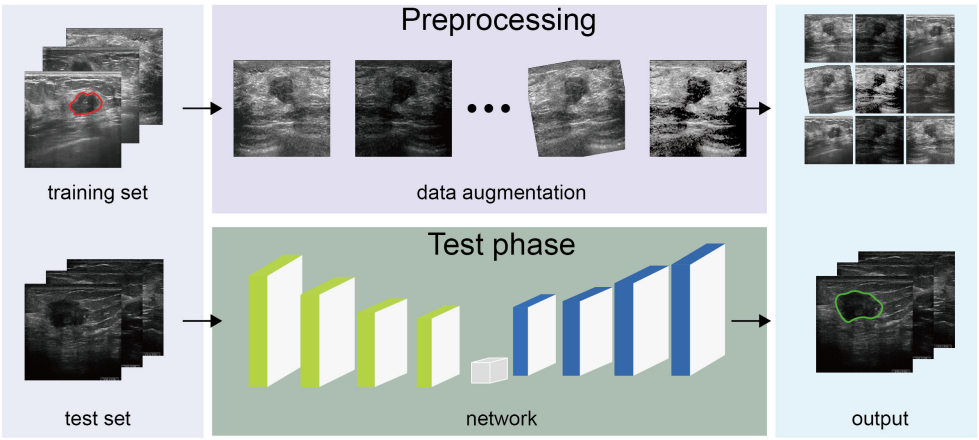


FIGURE 2
Algorithm construction Construction of the deep learning model.

calculate the refined boundary. After applying the above techniques and changing the network, our model could precisely segment breast tumors in ultrasound images.

Model training and model performance evaluation metrics

All datasets were collected from breast tissues examined in the ultrasound department. The dataset contains 1393 ultrasound (US) images. 956 US images of 856 patients were used as a training set to train our model, covering a variety of US images of benign breast tumors, malignant breast tumors, and post-chemotherapy breast cancer. And 437 NAC US images of 57 patients with complete NAC cycles were used as a test set to display network ability. There was no data overlap between the training set and the test set. We employed two experts to diagnose the whole dataset and generated corresponding ground truth masks. In detail, one expert annotated each image, while the other one reviewed it. When differences in diagnosis were encountered, the final annotation utilized the latter expert's results. Meanwhile, due to variation in US image size, during network training, we rescaled the ultrasound image with a fixed resolution of 448×384 pixels. Besides, in order to quickly find the optimizer weights, we adopted the Adam algorithm (39) with the betas from 0.5 to 0.999. When conducting experiments, we set the batch size 4 learning rate of 0.002 and stopped the iteration when the model was updated for 100 epochs, saving the most accurate model parameters and avoiding overfitting. All experiments were conducted in PyTorch under an Ubuntu OS server with an Intel Xeon (R) CPU E5-2680 v4 @2.40 GHz, 40 GB of RAM, and an NVIDIA GeForce RTX 3090 Ti with 24 GB of VRAM to boost training processing. We achieved our network on the 64-bits operation system and constructed algorithm on the Pytorch 2.0.1 framework with CuDNN 11.8. The training processing took up to 20 hours and the test phase lasted 76 seconds. Finally, the well-trained model generated a breast tumor distribution map for each ultrasound image and quantitatively analyzed performance.

To demonstrate the segmentation performance of our model, we utilized five metrics to quantitatively analyze the prediction output by comparing areas of prediction and annotation: accuracy, intersection over union (IoU), precision, recall, and the F1. All the metric formulas are show in Formula 1, 2, 3, 4, 5, and the TP, TF, FP and FN of each formula demonstrate the true positive, true negative, false positive and false negative. Among these metrics, accuracy, precision, recall, and the F1 score reflect the model's ability to capture specific features. Specifically, accuracy showcases the correctness of pixel predictions, precision demonstrates the model's capability to predict positive samples accurately, recall reflects the model's ability to capture positive samples, and the F1 score provides a balanced measure considering both positive and negative samples. As for the IoU, it denotes the intersection between the predicted segmentation and the ground truth divided by the area of union, which is commonly used in segmentation tasks. Considering the definition of the IoU metric, the value belongs to 0–1, and the closer the value is to 1, the more similar the prediction to the ground truth. To automatically calculate tumor parameters, we generated geometric parameters from an ultrasound image. After processing the model, we acquired a prediction mask for the tumor region. We first annotated each image with a scale bar, which could assist in precisely calculating the geometric information. Subsequently, we calculated the tumor number, area, diameter, area ratio, and length-to-width ratio using the scale bar and prediction mask. Accordingly, the IoU metric supported the overall network performance, whereas geometric parameters indicated the breast mass condition of each breast ultrasound image.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{IOU} = \frac{TP}{TP + FP + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{5}$$

Statistical analysis

Statistical analyses were performed using R version 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria). Normally distributed data are presented as the mean ± standard deviation. Normality was tested using the Shapiro–Wilk normality test. For comparisons between two groups, Student's t-test was used for normally distributed data, and the Wilcoxon rank-sum test was used for non-normally distributed data. For categorical variables, the chi-square test was used to determine differences between groups. The Kruskal–Wallis method was used to compare multiple groups. By utilizing the percentage reduction of tumors relative to their initial state at a specific time point (cycle N), as the predictor variable, and considering whether pCR was achieved as the outcome indicator, we compared the predictive performance of different models by constructing receiver operating characteristic (ROC) curves (Figure 3). We created ROC curves for all patients after completing the entire treatment cycle using four measurement models (Figure 3A); after each treatment cycle using the DLM (Figure 3C); after each treatment cycle using the MSM (Figure 3D); after each treatment cycle using the LAM (Figure 3E); and after each treatment cycle using the DAM (Figure 3F). It is essential to emphasize that Figure 3A includes patients who completed both 6 and 8 cycles of NAC. Therefore, the ROC curve in Figure 3A does not overlap with any ROC curves in Figures 3C–F. The p values of the area under the curve (AUC) were calculated using the DeLong test. A p-value<0.05 was considered statistically significant.

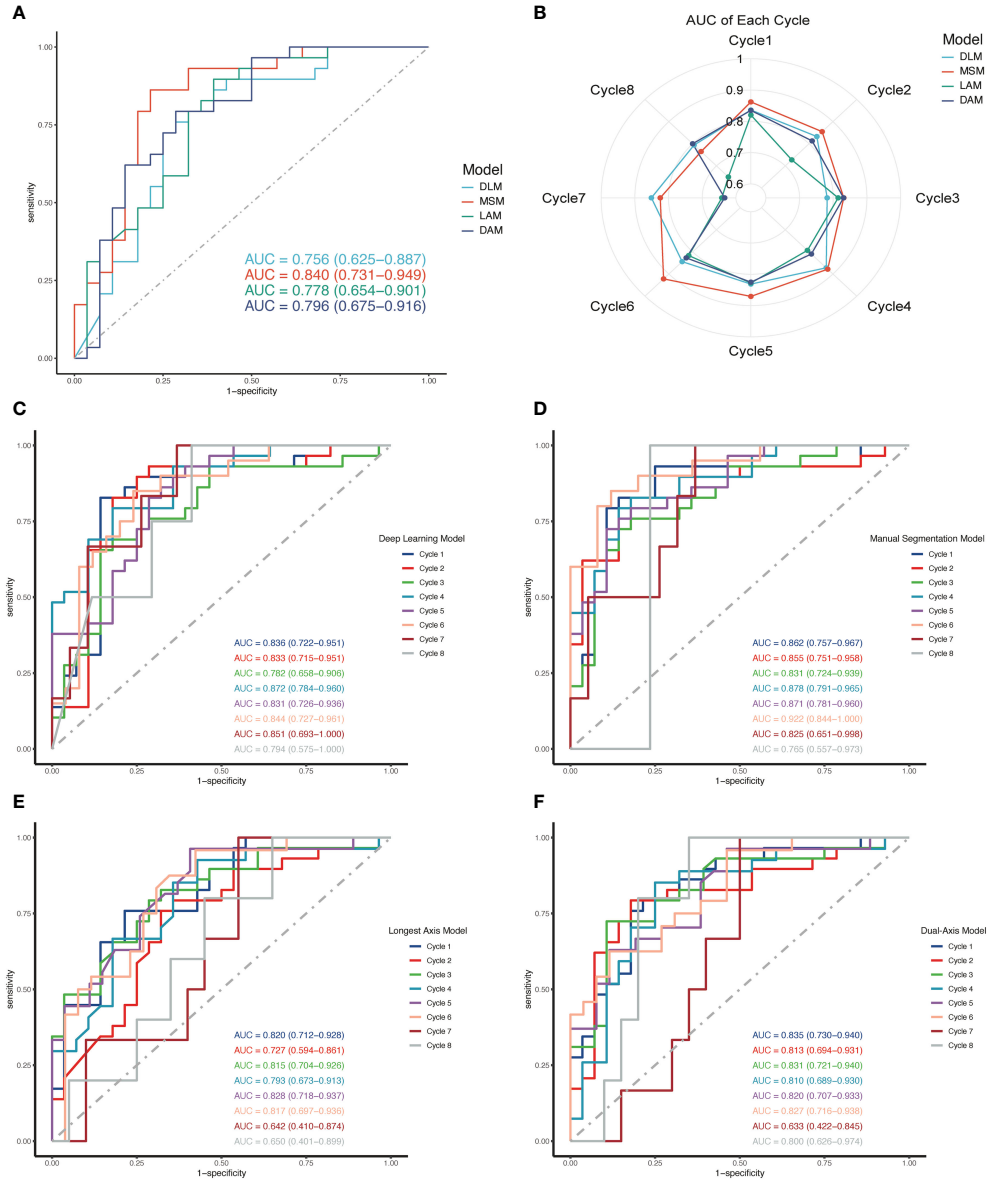


FIGURE 3
ROC curve for different models in predicting pCR. **(A)** The ROC curves were plotted to compare the ability of four different models in predicting pCR. **(B)** A radar chart shows the predictive ability of the four models at each cycle of the NAC. **(C–F)** After each cycle of NAC, tumor size was measured using four models and ROC curves were developed to predict pCR. The ROC curves for all cycles of the same model are drawn in one graph.

Results

Patient information

We included 57 patients who underwent complete per-cycle ultrasound assessment for primary invasive breast cancer. Table 1 summarizes basic patient data. The mean patient age was 49 years (range 28–70 years). Mean tumor size prior to NAC was calculated as follows: LAM 3.059 ± 1.283 cm, dual-axis model (DAM) 6.198 ± 5.734 cm^2, manual segmentation model (MSM) 4.190 ± 3.330 cm^2, and DLM 4.233 ± 3.638 cm^2. Considering all patients with breast cancer, 29 (50.9%) achieved pCR after NAC and 28 (49.1%) failed to achieve pCR. The pCR and non-pCR groups differed significantly in age (p=0.020), HER2 status (p<0.001), and post-NAC tumor size (LAM, p=0.030; DAM, p=0.008; MSM, p=0.031; DLM, p=0.009).

Performance of the DL-based model

After predicting the test-set images, we achieved a mean IoU of 0.856. Meantime, the model achieved best results on the dataset from the NAC, with an average accuracy of 0.973, average recall of 0.912 and average F1 score of 0. 918. Additionally, the segmentation capabilities of this model were quantitatively represented through ROC curve and PR curve (Supplementary Figure S1). The AUC for the ROC curve reached 0.99, and for the PR curve, it reached 0.92.

TABLE 1 Basic patient data.

Characteristics	Patients (n=57)		p-value
	pCR (n=29)	Non-pCR (n=28)	
Age (mean ± SD) (years)	52 ± 11	47 ± 11	0.020*
Histologic type			0.862
Invasive ductal carcinoma	18	18	
Others	11	10	
Clinical N stage			0.060
cN0	9	3	
cN1-3	20	25	
HER2 status			<0.001*
HER2+	21	4	
HER2-	8	24	
Pre-NAC tumor size (mean ± SD)			
LAM (cm)	3.321 ± 1.399	2.778 ± 1.103	0.114
DAM (cm^2)	7.133 ± 6.451	5.194 ± 4.765	0.209
MSM (cm^2)	4.936 ± 4.261	3.479 ± 2.704	0.112
DLM (cm^2)	4.874 ± 3.974	3.454 ± 2.316	0.135
Post-NAC tumor size (mean ± SD)			
LAM (cm)	1.148 ± 0.663	1.607 ± 0.870	0.030*
DAM (cm^2)	0.781 ± 0.829	1.726 ± 1.647	0.008*
MSM (cm^2)	0.549 ± 0.575	1.189 ± 1.115	0.031*
DLM (cm^2)	0.504 ± 0.561	0.958 ± 0.936	0.009*

*p<0.05 was considered significant. NAC, neoadjuvant chemotherapy; HER2, human epidermal growth factor receptor 2; pCR, pathological complete response; LAM, the longest axis model; DAM, dual-axis model; MSM, manual segmentation model; DLM, deep learning model; SD, standard deviation.

The DLM demonstrated good discrimination ability in addressing challenges related to delineating the boundary of breast cancer ultrasound images (e.g., blurred boundary, irregular shape, uneven brightness, and interference of contrast). And the DLM could also extract breast cancer feature of ultrasound image. Simultaneously, we also tested the model's computational complexity to display our network with more information. Our network contains 51.6 M parameters and 197.15G FLOPs and could detect a US within 0.08 seconds. In summary, the discrimination ability of the DLM was satisfactory for benign breast tumors and early-stage cancers undergoing NAC, but relatively poor for images of late-stage NAC and pCR.

Ablation experiment

To deepen our understanding of the mechanisms of our model, we conducted an ablation experiment. Based on the functionality of the model, we divided the algorithm into three modules: U-Net, ASPP, and self-attention. We designated U-Net as the baseline, named the model with combined multi-scale ASPP as baseline-M, and the U-Net with combined self-attention module as baseline-A. All experiments were conducted using the same ultrasound images and identical data preprocessing methods to analyze the effects of each module and the overall experimental performance. Additionally, we selected four ultrasound images with different manifestations for visual analysis, providing a comprehensive examination of the algorithm's performance from quantitative to qualitative perspectives.

In terms of cancer segmentation, the baseline model exhibited the weakest performance, with an average IoU index of only 0.742. The baseline-M, incorporating a multi-scale ASPP module, demonstrated superior performance by capturing ultrasound information at multiple scales and avoiding interference from noise or shadow areas. However, baseline-A, which integrated a self-attention mechanism into the baseline model, showed improved performance by directing more computational resources to cancer regions. Through comparative analysis, the performance gain of the baseline-A model was not as high as that of baseline-M. Nevertheless, the final model output results indicate that modules combining multi-scale and self-attention mechanisms can complement each other's shortcomings, achieving optimal performance. Finally, the model achieved the best results on the NAC dataset, with an average accuracy of 0.973, average recall of 0.912, average IoU of 0.856, and average F1 score of 0.918 (Table 2).

Simultaneously, we visually demonstrated the performance of each model. We selected ultrasound images from four NAC treatment cycles, encompassing both pCR and non-pCR response, as well as variations in tumor boundary clarity. In Figure 4, we illustrated that the clearer the tumor boundary, the better the image quality, the better the model's performance. However, in comparison to other models, our model resisted noise and shadows, accurately distinguishing the boundaries of cancer. In Figure 4, the baseline model exhibited the weakest resistance to interference, easily influenced by shadows and tissue around the tumor, especially in the second and third rows of the baseline images. Baseline-M, introducing a multi-scale features mechanism, significantly reduced interference from shadows and noise in ultrasound images. However, its ability to extract global information remains limited, and the accuracy of edge delineation is not precise. In the visual results of baseline-A, we observe that the attention mechanism consumes more computational resources on suspicious areas, but its detection ability is insufficient. In contrast, our proposed method, incorporating both multi-scale and self-attention mechanisms, enhances the segmentation ability to

10.3389/fonc.2024.1255618

TABLE 2 Metric results of different components on our clinical dataset from NAC.

	U-Net	ASPP	Attention	Class	Classification				Average				
					IoU	Precision	Recall	F1	Accuracy	Precision	Recall	IoU	F1
Baseline	√			**Normal**	0.952	0.958	0.993	0.975	0.955	0.926	0.78	0.742	0.835
				Cancer	0.531	**0.894**	0.567	0.695					
Baseline-M	√	√		**Normal**	0.970	0.978	0.991	0.985	0.972	**0.936**	0.886	0.842	0.909
				Cancer	0.714	0.893	0.781	0.833					
Baseline-A	√		√	**Normal**	0.962	0.977	0.984	0.981	0.965	0.903	0.875	0.812	0.888
				Cancer	0.661	0.829	0.766	0.796					
Ours	√	√	√	**Normal**	0.971	0.984	0.987	0.985	**0.973**	0.926	**0.912**	**0.856**	**0.918**
				Cancer	**0.740**	0.866	**0.836**	**0.851**					

Baseline: U-Net; Baseline-M: U-Net + Multi-scale ASPP; Baseline-A: U-Net + Self-Attention Module.

capture global information. As a result, we ultimately achieve robust cancer segmentation with improved resistance to interference.

Measures of tumor size ratios

Following NAC, both pCR and non-pCR groups exhibited tumor shrinkage. However, tumor regression was more significant in the pCR group than in the non-pCR group ($p<0.01$), particularly during the early stages of treatment (Figures 5B, C). As shown in Figure 5A, after the first two NAC cycles, the average residual tumor size measured by MSM, DLM, and DAM was <50% for patients in the pCR group, whereas the residual tumor size of patients in the non-pCR group was >50% (detailed data are displayed in the Supplementary Files).

Performance of the four models for predicting pCR

ROC curves were established for the four models to predict the possibility of obtaining pCR after the NAC completion in all patients (Figure 3A). The AUCs for LAM, DAM, DLM, and MSM were 0.778 (95% confidence interval [CI], 0.654–0.901), 0.796 (95% CI, 0.675–0.916), 0.756 (95% CI, 0.625–0.887), and 0.840 (95% CI, 0.731–0.949), respectively. The AUCs were compared using the DeLong test (DLM vs. LAM, $p=0.769$; DLM vs. DAM, $p=0.769$; DLM vs. MSM, $p=0.133$; LAM vs. DAM, $p=0.557$; LAM vs. MSM, $p=0.269$; and DAM vs. MSM, $p=0.358$), with no significance detected ($p<0.05$).

Subsequently, we plotted the ROC curves for each measurement method separately across the entire course of NAC and calculated the AUC to determine which model was more powerful in predicting pCR at an early stage (Figures 3B–F). After the DeLong's test, there was no significant difference between the AUCs of the different cycles ($p>0.05$), indicating that there was no significant difference in the predictive effect of the percentage of residual tumor size on pCR in the early (cycle 1-2), middle (cycle 3-4) and late (after 4 cycles) stages of NAC treatment. The AUC values of DLM for pCR prediction during all-phases of NAC treatment are as follows: cycle 1 (C1): 0.826; C2: 0.833; C3: 0.782; C4:0.872; C5:0.831; C6:0.844; C7: 0.851; C8:0.794 ($p>0.05$).

Discussion

In this study, we collected ultrasound images of breast cancer patients underwent NAC and developed an efficient DLM for tumor detection. In comparison to various methods employed for tumor segmentation in other studies, such as He et al.'s (27) utilization of the HCT-network, Xu et al.'s (40) implementation of region attention, Lyu et al.'s (41) application of the pyramid attention, and Chen et al.'s (42) use of the cascade network, our approach stands out by efficiently extracting global and local features using the WSA module and enhancing robustness using ASSP and FAM. Our model reaches the highest IOU of 0.856 among these studies

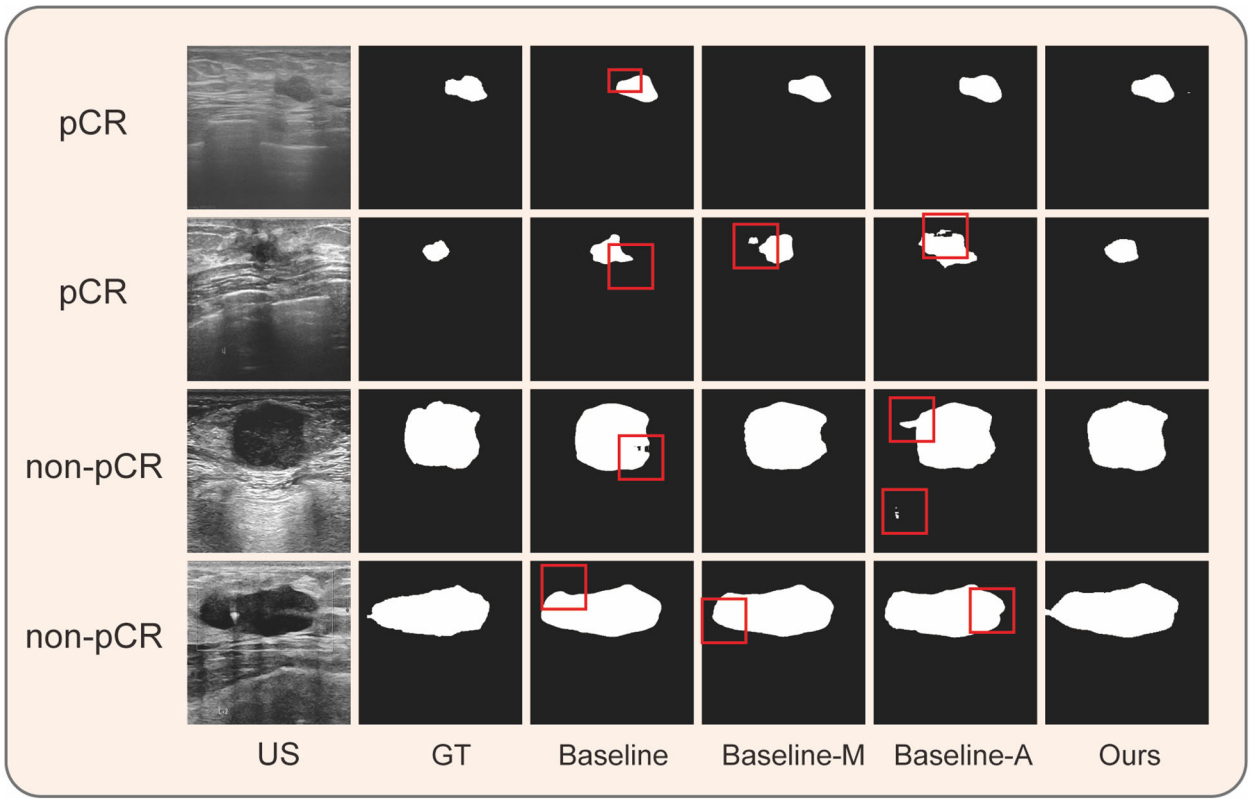


FIGURE 4
Performance of different components on four clinical ultrasound images during NAC. The results of ablation experiment. Baseline: U-Net; Baseline-M: U-Net + Multi-scale ASPP; Baseline-A: U-Net + Self-Attention Module; US, ultrasound; GT, ground truth.

while successfully distinguishing normal tissues and mitigating noise and shadow interference. Additionally, our dataset was both extensive and diverse which collected from clinical patients with a variety of characteristics displaying the comprehensiveness of breast ultrasound images, such as age, cancer grading, and pCR status. moreover, the data size consisting of 913 patients and 1393 images, significantly exceeds that of existing public datasets, such as the 780 images in BUSI (43) and the 163 images in DatasetB (44), ensuring suitability for breast ultrasound-specific tasks. Therefore, our DLM achieving excellent segmentation performance. Herein, we employed the DLM to measure changes in tumor size during NAC and predict pathological outcomes accordingly. We aimed to assess the capability of DLM for predicting breast cancer NAC outcomes.

Firstly, we compared the sensitivity of pCR and non-pCR breast cancers to NAC by constructing tumor size change ratios and line graphs. The percentage of tumor regression in patients with and without pCR was the most notably distinguished after the first NAC cycle, and the difference in the percentage of regression between the two groups gradually decreased. Ultrasound images and DLMs have been shown to predict the treatment response in the early (cycle 1–2) or mid-treatment (cycle 4) stages of NAC (32, 33, 45). However, follow-up observations of responses to all consecutive courses of NAC are still lacking. Hence, we further explored the accuracy of predicting pCR based on the percentage of tumor regression at different treatment stages of NAC. Based on our findings, there was no significant difference in the predictive ability of ultrasound assessment for pCR when performed at any NAC cycle; therefore, ultrasound assessment can be performed at any time, regardless of the treatment course.

Furthermore, we compared the predictive abilities of the four models for pCR after NAC for breast cancer (Figure 3). Tumor size is known to be closely related to the therapeutic effects of NAC (11). Conventional ultrasound models such as LAM and DAM assess tumor size by measuring dimensions. MSM and DLM assess tumor size by measuring area. The constructed DLM possesses the same level of capability as experienced sonographers in accurately identifying tumor boundaries. There was no significant difference in the predictive efficacy of the four models for pCR. Therefore, the DLM can serve as a valuable tool for assisting sonographers in manually measuring breast cancer, enabling more precise calculations of tumor size. Notably, our DLM is still helpful in clinical practice. First, it helps improve the workflow. The algorithms rapidly and automatically identify tumor areas within ultrasound images, allowing for precise segmentation and accurate tumor size measurements. The DLM developed in this study offers a time-efficient model to reduce the burden on the sonographers and eliminate bias due to differences in experience. The predictive effect of the DLM is also relatively stable throughout the treatment process compared to conventional measurement models (Figure 3B). Second, our DLM was learned from the database of a large general hospital and reviewed by senior sonographers, and the

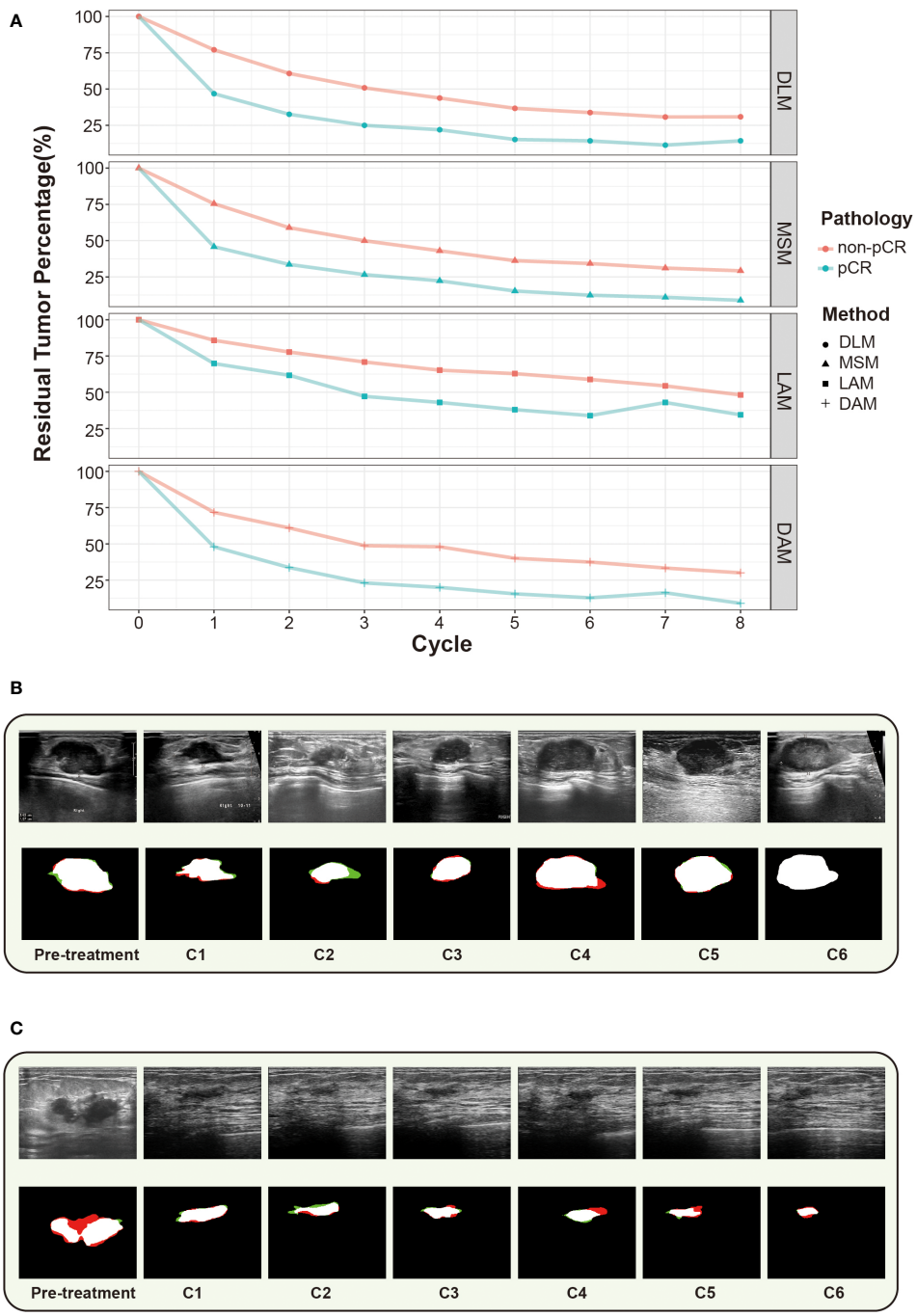


FIGURE 5
Changes in residual tumor with NAC cycles in the pCR and non-pCR groups. **(A)** The percentage of residual tumors decreased with the increase of NAC cycle in both groups, with the most evident decrease in the first two cycles. **(B)** in ultrasound images of breast cancer of a patient who didn't receive pCR. **(C)** Changes in ultrasound images of breast cancer of a patient who received pCR. The green line denotes the prediction contour, while the red line denotes the ground truth in the first column.

DLM's ability to recognize ultrasound images is equivalent to that of an experienced sonographer. Therefore, the DLM is useful for assisting junior sonographers or primary hospitals in breast ultrasound examinations. Third, our research helps doctors and patients by facilitating the tracking of tumor growth and treatment responses. It enables the recording of tumor size and shape at different time points, allowing for image comparisons. This automatic cancer detection capability enables clinicians to evaluate treatment effectiveness and adjust treatment plans, as necessary. Most importantly, predicting tumor responses and tailor personalized treatment plans to ensure the best treatment outcomes for patients.

The direct comparison of our results with those of other studies can be challenging, given the differences in data collection and analysis methods. To the best of our knowledge, this is the first study to use DL to continuously monitor changes in tumor size

TABLE 3 Similar studies.

Studies	AUC
Candelaria et al.	0.760 for patients with TNBC 0.822 for HR+/HER2- patients 0.522 for HR-/HER2+ patients
Gounaris et al.	0.689
Byra et al.	0.797
Sannachi et al.	0.780-0.860

AUC, area under the curve; TNBC, triple-negative breast cancer; HER2, human epidermal growth factor receptor 2.

during NAC. The previous studies similar to this research and their AUC values for predicting pCR have been listed in Table 3. Candelaria et al. (45) and Gounaris et al. (46) predicted pCR by measuring the largest change in tumor diameter at mid-treatment. Sannachi et al. (47) employed quantitative ultrasonography to collect data at the 1st, 4th, and 8th NAC cycles. Byra et al. (32) used a DL approach based on early treatment ultrasound images to assess the treatment response. The authors also performed consecutive predictions during the first four cycles, concluding the absence of any significant difference in the prediction of chemotherapy outcomes during the first four cycles (48), consistent with our findings.

In the present study, the following innovations are pivotal. First, the DL method for processing ultrasound images has several advantages. 1) Our DL method based on U-Net is an automatic end-to-end neural network that does not manually screen handcrafted features. The model can infer the entire ultrasound image and provide prediction results directly. Moreover, the model can distinguish the tumor boundary from breast tissue by intelligently extracting morphological features from cancer and normal tissues. Meanwhile, the prediction image had the same dimensions as the input image, displaying breast cancer distribution. 2) Our model adopted a multi-scale method and attention mechanism to improve the ability to extract and focus on features for accurate and rapid tumor tissue identification. Moreover, the multi-scale method considers context information to improve boundary performance, particularly the relationship between the tumor and normal tissue region. 3) To enhance the detection capacity, we adopted a hybrid loss function to search for optimal model parameters. By improving the accuracy of the tumor contour and pixel segmentation, our hybrid loss function consists of binary cross-entropy loss, dice loss, and active contour models. Second, to the best of our knowledge, this is the first study to evaluate ultrasound findings during NAC. Herein, the ultrasound assessment values were similar for each cycle. Third, we compared the effectiveness of conventional ultrasound tumor measurements and DLM in predicting pathologic response, with the DLM as an alternative to manual measurements.

Nevertheless, the limitations of the present study need to be addressed. First, this was a single-center study. Although this study was conducted over 2 years, recruiting an adequate number of representative patients was challenging. Further external validation should be performed by recruiting more patients in a multicenter prospective study to demonstrate the accuracy of the DLM. Second, the DLM predictions were solely based on changes in tumor size. Although size change is the most important and intuitive indicator of tumors' response to chemotherapy, this approach overlooks other ultrasound image features related to the response to NAC, such as texture changes, tumor shrinkage patterns, and lymph node status. Including these additional features may further enhance the predictive performance of the model. Third, although ultrasound is one of the most important imaging modalities for breast cancer NAC, predictions based on ultrasound images alone are insufficient to determine whether the NAC treatment regimen should be stopped or changed. Many guidelines suggest employing dynamic contrast enhanced (DCE) MRI for assessing the efficacy of NAC. However, this study lacks a comparative analysis between ultrasound and MRI. Therefore, the clinical value of the DLM based on breast ultrasound for pCR outcome prediction remains to be validated.

Future research will be primarily focused on several key areas. First, there should be a focus on the multimodal fusion of medical imaging data to provide a more comprehensive medical insight for predicting the tumor's response. Second, there is a need for development of automated annotation tools to alleviate the burden of data labeling for medical professionals and researchers. Third, there should be a quantification of uncertainty and research into interpretability, particularly in medical ultrasound imaging. Fourth, there should be a focus on incremental learning to enable continuous adaptation to new data, with a special emphasis on monitoring long-term tumor changes and addressing new tumor types. These research directions will help enhance the efficiency and accuracy of medical image analysis, fostering ongoing improvements in tumor diagnosis and treatment.

Conclusions

In conclusion, we constructed a U-Net-based end-to-end DLM for processing ultrasound images during NAC treatment for breast cancer, which offers advantages of accuracy, efficiency, and automation to assist manual measurement. This study provides a non-invasive method for predicting individualized responses in breast cancer patients undergoing NAC at all stages of treatment.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Ethics Committee of Renmin Hospital of Wuhan University (approval number: WDRY2022-K217). The studies were conducted in accordance with the local legislation and institutional

requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JZ: Resources, Writing – original draft. JD: Formal Analysis, Writing – original draft. JH: Formal Analysis, Writing – original draft. LM: Methodology, Writing – review & editing. NL: Resources, Writing – review & editing. FY: Supervision, Writing – review & editing. CL: Methodology, Project administration, Writing – review & editing. SS: Supervision, Writing – review & editing. YZ: Funding acquisition, Project administration, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported in part by the National Natural Science Foundation of China (81502665); National Key R&D Program of China (2023YFF0723300); Science Fund for Distinguished Young Scholars of Hubei Province (2021CFA042); Hubei Province Young Science and Technology Talent Morning Hight Lift Project (202319); Jiangsu Science and Technology Program (BK20221257); Shenzhen Science and Technology Program (JCYJ20220530140601003, JCYJ20230807090207014) and Interdisciplinary Innovative Talents Foundation from Renmin Hospital of Wuhan University(JCRCYR-2022-014). The funder was used for technical supports, software purchases and text editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1255618/full#supplementary-material

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin* (2023) 73(1):17–48. doi: 10.3322/caac.21763

2. Giordano SH. Update on locally advanced breast cancer. *Oncologist* (2003) 8 (6):521–30. doi: 10.1634/theoncologist.8-6-521

3. Charfare H, Limongelli S, Purushotham AD. Neoadjuvant chemotherapy in breast cancer. *Vol. 92 Br J Surgery.* (2005) 92:14–23. doi: 10.1002/bjs.4840

4. Zardavas D, Piccart M. Neoadjuvant therapy for breast cancer. *Annu Rev Med* (2015) 66:31–48. doi: 10.1146/annurev-med-051413-024741

5. Gradishar WJ, Moran MS, Abraham J, Aft R, Agnese D, Allison KH, et al. Breast cancer, version 3.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Cancer Network.* (2022) 18:452–478. doi: 10.6004/jnccn.2020.0016

6. CACA-CBCS CACAC of BCS. Chinese anti-cancer society breast cancer diagnosis and treatment guidelines and standards. *Chinaoncology* (2021) 31(8):770–856.

7. Viale G, Fusco N. Pathology after neoadjuvant treatment – How to assess residual disease. *Breast* (2022) 62:S25–8. doi: 10.1016/j.breast.2021.11.009

8. Symmans WF, Peintinger F, Hatzis C, Rajan R, Kuerer H, Valero V, et al. Measurement of residual breast cancer burden to predict survival after neoadjuvant chemotherapy. *J Clin Oncol* (2007) 25(28):4414–22. doi: 10.1200/JCO.2007.10.6823

9. Spring L, Greenup R, Niemierko A, Schapira L, Haddad S, Jimenez R, et al. Pathologic complete response after neoadjuvant chemotherapy and long-term outcomes among young women with breast cancer. *JNCCN J Natl Compr Cancer Network.* (2017) 15(10):1216–23. doi: 10.6004/jnccn.2017.0158

10. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: The CTNeoBC pooled analysis. *Lancet* (2014) 384(9938):164–72. doi: 10.1016/S0140-6736 (13)62422-8

11. Schott AF, Roubidoux MA, Helvie MA, Hayes DF, Kleer CG, Newman LA, et al. Clinical and radiologic assessments to predict breast cancer pathologic complete response to neoadjuvant chemotherapy. *Breast Cancer Res Treat* (2005) 92(3):231–8. doi: 10.1007/s10549-005-2510-1

12. Dialani V, Chadashvili T, Slanetz PJ. Role of imaging in neoadjuvant therapy for breast cancer. *Ann Surg Oncol Springer New York LLC;* (2015) 22:1416–24. doi: 10.1245/s10434-015-4403-9

13. Park SH, Moon WK, Cho N, Song IC, Chang JM, Park IA, et al. Diffusion-weighted MR imaging: Pretreatment prediction of response to neoadjuvant chemotherapy in patients with breast cancer. *Radiology* (2010) 257(1):56–63. doi: 10.1148/radiol.10092021

14. Gu YL, Pan SM, Ren J, Yang ZX, Jiang GQ. Role of magnetic resonance imaging in detection of pathologic complete remission in breast cancer patients treated with neoadjuvant chemotherapy: A meta-analysis. *Clin Breast Cancer. Elsevier Inc.* (2017) 17:245–55. doi: 10.1016/j.clbc.2016.12.010

15. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev BioMed Eng* (2017) 19(1):221–48. doi: 10.1146/annurev-bioeng-071516-044442

16. Lei YM, Yin M, Yu MH, Yu J, Zeng SE, Lv WZ, et al. Artificial intelligence in medical imaging of the breast. *Front Oncol* (2021) 11. doi: 10.3389/fonc.2021.600557

17. Freeman K, Geppert J, Stinton C, Todkill D, Johnson S, Clarke A, et al. Use of artificial intelligence for image analysis in breast cancer screening programmes: Systematic review of test accuracy. *BMJ.* (2021) 374:n1872. doi: 10.1136/bmj.n1872

18. Stember JN, Terilli KL, Perez E, Megjhani M, Cooper CA, Jambawalikar S, et al. Surface point cloud ultrasound with transcranial doppler: coregistration of surface point cloud ultrasound with magnetic resonance angiography for improved reproducibility, visualization, and navigation in transcranial doppler ultrasound. *Journal of Digital Imaging* (2020) 33:930–936. doi: 10.1007/s10278-020-00328-y

19. Jiang M, Li CL, Luo XM, Chuan ZR, Lv WZ, Li X, et al. Ultrasound-based deep learning radiomics in the assessment of pathological complete response to neoadjuvant chemotherapy in locally advanced breast cancer. *Eur J Cancer.* (2021) 147:95–105. doi: 10.1016/j.ejca.2021.01.028

20. Ilesanmi AE, Idowu OP, Chaumrattanakul U, Makhanov SS. Multiscale hybrid algorithm for pre-processing of ultrasound images. *BioMed Signal Process Control.* (2021) 66:102396. doi: 10.1016/j.bspc.2020.102396

21. Balkenende L, Teuwen J, Mann RM. Application of deep learning in breast cancer imaging. *Semin Nucl Med W.B. Saunders;* (2022) 52:584–96. doi: 10.1053/j.semnuclmed.2022.02.003

22. Lei YM, Yin M, Yu MH, Yu J, Zeng SE, Lv WZ, et al. Artificial intelligence in medical imaging of the breast. *Front Oncol* (2021) 11. doi: 10.3389/fonc.2021.600557

23. Al-Dhabyani W, Khaled H, Fahmy A, Gomaa M. Deep Learning Approaches for Data Augmentation and Classification of Breast Masses using Ultrasound Images Ultrasound Breast Diagnosis Using Machine Learning View project Deep Learning Approaches for Data Augmentation and Classification of Breast Masses using Ultrasound Images. *Article Int J Advanced Comput Sci Appl* (2019) 10(5):618–627. doi: 10.14569/IJACSA.2019.0100579

24. Chen G, Liu Y, Qian J, Zhang J, Yin X, Cui L, et al. DSEU-net: A novel deep supervision SEU-net for medical ultrasound image segmentation. *Expert Syst Appl* (2023) 223:119939. doi: 10.1016/j.eswa.2023.119939

25. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*. Boston, MA, USA: IEEE. (2015). 3431–40. doi: 10.1109/CVPR.2015.7298965

26. Liang Y, He R, Li Y, Wang Z. Simultaneous segmentation and classification of breast lesions from ultrasound images using Mask R-CNN, in: *2019 IEEE International Ultrasonics Symposium (IUS)*. Glasgow, UK: IEEE. (2019). 1470–2. doi: 10.1109/ULTSYM.2019.8926185

27. He Q, Yang Q, Xie M. HCTNet: A hybrid CNN-transformer network for breast ultrasound image segmentation. *Comput Biol Med* (2023) 155:106629. doi: 10.1016/j.compbiomed.2023.106629

28. Zhang J, Liu Y, Wu Q, Wang Y, Liu Y, Xu X, et al. SWTRU: Star-shaped Window Transformer Reinforced U-Net for medical image segmentation. *Comput Biol Med* (2022) 150:105954. doi: 10.1016/j.compbiomed.2022.105954

29. Lin G, Chen M, Tan M, Chen L, Chen J. A dual-stage transformer and MLP-based network for breast ultrasound image segmentation. *Biocybern BioMed Eng.* (2023) 43(4):656–71. doi: 10.1016/j.bbe.2023.09.001

30. Yuan F, Zhang Z, Fang Z. An effective CNN and Transformer complementary network for medical image segmentation. *Pattern Recognit* (2023) 136:109228. doi: 10.1016/j.patcog.2022.109228

31. Peng J, Pu H, Jia Y, Chen C, Ke XK, Zhou Q. Early prediction of response to neoadjuvant chemotherapy using contrast-enhanced ultrasound in breast cancer. *Medicine* (2021) 100(19):e25908. doi: 10.1097/MD.0000000000025908

32. Byra M, Dobruch-Sobczak K, Klimonda Z, Piotrzkowska-Wroblewska H, Litniewski J. Early prediction of response to neoadjuvant chemotherapy in breast cancer sonography using siamese convolutional neural networks. *IEEE J BioMed Health Inform.* (2021) 25(3):797–805. doi: 10.1109/JBHI.2020.3008040

33. Gu J, Tong T, He C, Xu M, Yang X, Tian J, et al. Deep learning radiomics of ultrasonography can predict response to neoadjuvant chemotherapy in breast cancer at an early stage of treatment: a prospective study. *Eur Radiol* (2022) 32(3):2099–109. doi: 10.1007/s00330-021-08293-y

34. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer.* (2009) 45(2):228–47. doi: 10.1016/j.ejca.2008.10.026

35. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* (2013) 31(31):3997–4013. doi: 10.1200/JCO.2013.50.9984

36. Yan P, Sun Q, Yin N, Hua L, Shang S, Zhang C. Detection of coal and gangue based on improved YOLOv5.1 which embedded scSE module. *Measurement* (2022) 188:110530. doi: 10.1016/j.measurement.2021.110530

37. Chen X, Williams BM, Vallabhaneni SR, Czanner G, Williams R, Zheng Y. Learning active contour models for medical image segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE (2019). 11624–32. doi: 10.1109/CVPR.2019.01190

38. Karimi D, Salcudean SE. Reducing the hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Trans Med Imaging.* (2020) 39 (2):499–513. doi: 10.1109/TMI.2019.2930068

39. Kingma DP, Lei Ba J. Adam: A method for stochastic optimization. *International Conference on Learning Representations*. (2015). doi: 10.48550/arXiv.1412.6980

40. Xu M, Huang K, Qi X. A regional-attentive multi-task learning framework for breast ultrasound image segmentation and classification. *IEEE Access* (2023) 11:5377–5392. doi: 10.1109/ACCESS.2023.3236693

41. Lyu Y, Xu Y, Jiang X, Liu J, Zhao X, Zhu X. AMS-PAN: Breast ultrasound image segmentation model combining attention mechanism and multi-scale features. *BioMed Signal Process Control.* (2023) 81:104425. doi: 10.1016/j.bspc.2022.104425

42. Chen G, Dai Y, Zhang J. C-Net: Cascaded convolutional neural network with global guidance and refinement residuals for breast ultrasound images segmentation. *Comput Methods Programs Biomed* (2022) 225:107086. doi: 10.1016/j.cmpb.2022.107086

43. Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. *Data Brief.* (2020) 28:104863. doi: 10.1016/j.dib.2019.104863

44. Huang Q, Huang Y, Luo Y, Yuan F, Li X. Segmentation of breast ultrasound image with semantic classification of superpixels. *Med Image Anal* (2020) 61:101657. doi: 10.1016/j.media.2020.101657

45. Candelaria RP, Bassett RL, Symmans WF, Ramineni M, Moulder SL, Kuerer HM, et al. Performance of mid-treatment breast ultrasound and axillary ultrasound in predicting response to neoadjuvant chemotherapy by breast cancer subtype. *Oncologist [Internet].* (2017) 22(4):394–401. doi: 10.1634/theoncologist.2016-0307

46. Gounaris I, Provenzano E, Vallier AL, Hiller L, Iddawela M, Hilborne S, et al. Accuracy of unidimensional and volumetric ultrasound measurements in predicting good pathological response to neoadjuvant chemotherapy in breast cancer patients. *Breast Cancer Res Treat [Internet].* (2011) 127(2):459–69. doi: 10.1007/s10549-011-1454-x

47. Sannachi L, Gangeh M, Tadayyon H, Sadeghi-Naini A, Gandhi S, Wright FC, et al. Response monitoring of breast cancer patients receiving neoadjuvant chemotherapy using quantitative ultrasound, texture, and molecular features. *PloS One* (2018) 13(1). doi: 10.1371/journal.pone.0189634

48. Byra M, Dobruch-Sobczak K, Piotrzkowska-Wroblewska H, Klimonda Z, Litniewski J. Prediction of response to neoadjuvant chemotherapy in breast cancer with recurrent neural networks and raw ultrasound signals. *Phys Med Biol* (2022) 67 (18). doi: 10.1088/1361-6560/ac8c82

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 29 February 2024
DOI 10.3389/fonc.2024.1320714

OPEN ACCESS

EDITED BY
Jiabao He,
Newcastle University, United Kingdom

REVIEWED BY
Fenglin Dong,
The First Affiliated Hospital of Soochow University, China
Bilgin Kadri Aribas,
Bülent Ecevit University, Türkiye

*CORRESPONDENCE
Qiao Hu
xyyxyhq@163.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 14 October 2023
ACCEPTED 13 February 2024
PUBLISHED 29 February 2024

CITATION
Li L-L, Su Q-L, Deng Y-X, Guo W-W, Lun H-M and Hu Q (2024) Contrast-enhanced ultrasound for the preoperative prediction of pathological characteristics in breast cancer.
Front. Oncol. 14:1320714.
doi: 10.3389/fonc.2024.1320714

Contrast-enhanced ultrasound for the preoperative prediction of pathological characteristics in breast cancer

Ling-Ling Li[1†], Quan-Li Su[1†], Yun-Xia Deng[1], Wen-Wen Guo[2], Hai-Mei Lun[1] and Qiao Hu[1*]

[1]Departments of Ultrasound, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, China, [2]Departments of Pathology, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, China

Objective: We aimed to investigate the value of contrast-enhanced ultrasound (CEUS) in the preoperative prediction of the histological grades and molecular subtypes of breast cancer.

Methods: A total of 183 patients with pathologically confirmed breast cancer were included. Contrast enhancement patterns and quantitative parameters were compared in different groups. The receiver operating characteristic (ROC) curve was used to analyze the efficacy of CEUS in the preoperative prediction of pathological characteristics, including histologic grade and molecular subtypes.

Results: Heterogeneous enhancement, perfusion defects, and peripheral radial vessels were mostly observed in higher histologic grade (grade III) breast cancer. Heterogeneous enhancement and perfusion defect were the most effective indicators for grade III breast cancer, with the areas under the ROC curve of 0.768 and 0.756, respectively. There were significant differences in the enhancement intensity, post-enhanced margin, perfusion defects, and peripheral radial vessel among the different molecular subtypes of breast cancer (all $P < 0.01$). Perfusion defects and clear edge after enhancement were the best qualitative criteria for the diagnosis of HER-2 overexpressed and triple-negative breast cancers, and the corresponding areas under the ROC curves were 0.804 and 0.905, respectively. There were significant differences in PE, WiR, WiPI, and WiWoAUC between grade III vs grade I and II breast cancer ($P < 0.05$). PE, WiR, WiPI, and WiWoAUC had good efficiency in the diagnosis of high-histologic-grade breast cancer. PE had the highest diagnostic efficiency in Luminal A, while WiPI had the highest diagnostic efficiency in Luminal B subtype breast cancer, and the areas under the ROC curve were 0.825 and 0.838, respectively. WiWoAUC and WiR were the most accurate parameters for assessing triple-negative subtype breast cancers, and the areas under the curve were 0.932 and 0.922, respectively.

Conclusion: Qualitative and quantitative perfusion analysis of contrast-enhanced ultrasound may be useful in the non-invasive prediction of the histological grade and molecular subtypes of breast cancers.

KEYWORDS

contrast-enhanced ultrasound, VueBox, breast cancer, histological grade, molecular subtypes

Introduction

According to the latest analysis from the International Agency for Research on Cancer, new cases of breast cancer surpassed those of lung cancer in 2020, becoming the world's most commonly diagnosed cancer (1). Breast cancer is a highly heterogeneous tumor, and the pathological grading of breast cancer reflects the degree of malignancy and invasiveness of the tumor to a certain extent (2). The higher the pathological grading, the more malignant the tumor, and the worse the prognosis. The expression levels of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) in breast cancer cells are considered important factors in determining the biological behavior of breast cancer and the efficacy of endocrine therapy and are the main prognosis predictors (2, 3). For the preoperative assessment of the histopathological grading and molecular subtyping of breast cancer, the clinical standard is the pathological immunohistochemistry of needle biopsies. However, pathological biopsy is an invasive procedure and has the defect of insufficient samples to make a diagnosis. Therefore, it is necessary to develop non-invasive imaging modalities to predict the pathological characteristics of breast cancer before surgery.

VueBox is a color-coded external perfusion software program that can be used for dynamic contrast-enhanced ultrasound (CEUS) with motion/respiration compensation and is suitable for DICOM format video images acquired by various ultrasound equipment (4). The software calculates the perfusion parameters automatically, generating color-coded maps of the perfusion parameters and thus providing a more direct and objective quantitative analysis of the subtle difference in enhancement degree, reducing the subjective dependence of image interpretations by operators. In recent years, there have been literature reports on the application of VueBox to quantitatively analyze the characteristics of breast, thyroid, liver, and other organ tumors by CEUS (5–7). Jung EM (7) et al. used the TIC curve of VueBox external perfusion software to compare and analyze the perfusion performance of benign and malignant non-cystic breast masses. The results showed that the quantitative CEUS perfusion parameters PE and areas under the curve (AUC) can well evaluate the malignant risk of non-cystic breast masses. This may reduce the risk rating for certain BI-RADS category 4 lesions. However, there are few reports on the application of CEUS perfusion imaging with VueBox for the evaluation of breast cancer pathological characteristics. In this study, VueBox external perfusion analysis software was used to explore the value of contrast-enhanced ultrasound (CEUS) in the preoperative prediction of the pathological grading and molecular subtyping of breast cancer.

Materials and methods

Patients

This study was performed with the approval of the Ethics Committee of the People's Hospital of Guangxi Zhuang Autonomous Region, China (IRB No. KY-LW-2020-24). Informed consent was obtained from all participants. A total of 183 patients with breast cancer were enrolled from December 2020 to April 2023 (women, age range 28-85 y, and mean age 52 y). Inclusion criteria: ① patients who underwent CEUS and for whom contrast dynamic images were stored; ② no preoperative treatment; ③ pathologically confirmed invasive breast cancer; and ④ complete postoperative pathological and immunohistochemical results. The exclusion criteria were as follows: ① poor quality of CEUS dynamic images; and ② incomplete clinical data.

Imaging acquisition

Conventional ultrasound and CEUS imaging were performed using a GE LOGIQ E9 ultrasound system (GE Healthcare, Chicago, IL, USA) with a high-resolution linear transducer; the probe frequency was 6-15 MHz for conventional ultrasound and 6-9 MHz for CEUS. The patients were placed in a supine position with arms placed above the head. First, the whole breast was continuously multi-section scanned by conventional ultrasound. Color and power Doppler were performed in different planes to evaluate the intralesional vascularity. Then, the plane with the most abundant vessels including the lesion and its surrounding normal breast tissue was selected and switched to CEUS imaging mode with a mechanical index (MI) < 0.10. The ultrasound contrast agent used in the present study was SonoVue (Bracco SpA, Milan, Italy). A bolus of 4.8 mL of contrast agent was administrated via a peripheral vein and was immediately followed by a flush of 5 mL saline. CEUS continuous dynamic imaging was observed immediately after injection of the contrast agent for at least 6 minutes. The images and video clips (the last 2 minutes after contrast agent injection) were stored and transferred to a mobile hard disk in Dicom format for subsequent offline analysis.

CEUS image analysis

CEUS image analysis was performed by two radiologists (LL.L and Q.H with 3 and 12 years of experience in breast CEUS, respectively). They were blinded to the clinical data and pathological results of the patients. In cases of discrepancies, the two reviewers reanalyzed and discussed together to reach a consensus.

The following CEUS qualitative indicators were analyzed: (1) compared with that of surrounding normal breast tissue, the enhancement intensity was classified as hyper-, iso-, or hypo-enhancement; (2) based on the internal homogeneity of the tumor, enhancement was divided into homogeneous or heterogeneous enhancement; (3) the enhancement edge of the lesion was classified as a clear or blurred margin; (4) whether the lesion scope enlarged after enhancement; (5) the presence or absence of perfusion defect; (6) the presence or absence of radial or penetrating vessels.

Quantitative analysis for dynamic CEUS imaging was performed using the color-coded off-line software (VueBox, Bracco, Genève, Suisse). Regions of interest were manually delineated with the strongest enhanced area in the lesion and the

surrounding normal breast tissue at the same depth. Time-intensity curves were generated to obtain quantitative parameters, including peak enhancement (PE), mean transit time (mTT), time to peak (TTP), wash-in rate (WiR), wash-in perfusion index (WiPI), and wash-in and wash-out areas under the curve (WiWoAUC). For each quantitative data, the measurements were repeated three times, and their mean was used in the analysis.

Histopathologic analysis

Histopathological specimens were fixed with 10% formalin, embedded in paraffin, and sliced into 3-μm sections, and hematoxylin–eosin staining was performed. Invasive breast cancer was graded using the Nottingham histological grading system (8). Immunohistochemical staining was used to determine the expressions of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (Her-2), and Ki-67. The molecular classification of breast cancer was divided into four subtypes (9): ① luminal A subtype: ER(+) and/or PR(+), Her-2(-), and Ki-67 low expression (< 14%); ② luminal B subtype: ER(+) and/or PR(+), Her-2(+), or ER(+) and/or PR(+), Her-2(-), and Ki-67 high expression (≥ 14%); ③ Her-2 overexpressed subtype: ER(-), PR(-), and Her-2(+); and ④triple-negative subtype: ER(-), PR(-), and HER-2(-). All histopathologic slides were observed by a pathologist (WW.G) who had more than 6 years of experience in breast pathologic analysis.

Statistical analysis

SPSS 26.0 statistical software (IBM Corp, Armonk, NY, USA) was used for the statistical analysis. Count data are expressed as frequency (n), and intergroup comparisons were conducted using the χ^2 test. Quantitative parameters with a non-normal distribution were presented as M (Q1, Q3); the Kruskal–Wallis H test was used for intergroup comparisons of quantitative data, and the Dunn-Bonferroni test was used for pairwise comparisons. MedCalc software was also used to draw ROC curves for the parameters with significant differences to verify their diagnostic efficacy. $P < 0.05$ was considered statistically significant.

Results

There were 183 invasive breast cancers among 183 patients enrolled in this present study, including invasive ductal carcinoma (n = 152, 83.06%), invasive lobular carcinoma (n = 20, 10.93%), intraductal papillary carcinoma (n = 3, 1.64%), mucinous carcinoma (n = 3, 1.64%), and medullary carcinoma (n = 5, 2.73%). Of all the breast lesions (size, 28.8 ± 19.2 mm, range, 8.0-118mm), 126 (68.85%) were moderately (grade II) and highly differentiated (grade I), 57 (31.15%) were poorly differentiated (grade III), 50 (27.32%) were luminal A, 80 (43.72%) were luminal B, 31 (16.94%) were HER-2 overexpressed, and 22 (12.02%) were triple-negative subtypes. There were no significant differences in age and tumor size among patients with different grades and subtypes (Table 1).

Qualitative CEUS features of breast cancer with different pathological grades and molecular subtypes

The qualitative CEUS analysis revealed that higher histological grade (grade III) breast cancer mostly showed heterogeneous enhancement (50/57, 87.72%), perfusion defect (41/57, 71.93%), and presence of radial or penetrating vessels (47/57, 82.46%). Lower histological grade (grade I and II) breast cancer showed more iso- or hypo-enhancement (79/126, 62.70%), homogeneous enhancement (83/126, 65.87%), and no obvious perfusion defect (100/126, 79.37%) (Figure 1). The enhancement degree, internal homogeneity, perfusion defect, and presence or absence of radial or penetrating vessels showed significant differences between lower histological grade and higher histological breast cancer (all P <0.01). However, with regard to the enhancement edge and whether lesion scope enlarged after enhancement, no statistical difference was found between the two groups ($P > 0.05$) (Table 2).

Iso- or hypo-enhancement was found in 43 (43/50, 86.00%) luminal A and 41 (41/80, 51.25%) luminal B subtype breast cancers. There were 44 (44/50, 88.00%) luminal A and 42 (42/80, 52.25%) luminal B lesions present in the radial or penetrating vessels. The

TABLE 1 Clinicopathological Characteristics of patients.

Characteristics	Case number (percentage)
Age (years)*	52.21 ± 11.54
Tumor diameter (mm)*	28.82 ± 19.24
Histologic grade	
Grades I and II	126 (68.85%)
Grade III	57 (31.15%)
Molecular subtypes	
Luminal A	50 (27.32%)
Luminal B	80 (43.72%)
Her-2 overexpressed	31 (16.94%)
Triple-negative	22 (12.02%)
Histologic type	
Invasive ductal carcinoma	152 (83.06%)
Invasive lobular carcinoma	20 (10.93%)
Intraductal papillary carcinoma	3 (1.64%),
Mucinous carcinoma	3 (1.64%)
Medullary carcinoma	5 (2.73%)
Lymph node status	
Positive	60 (32.79%)
Negative	123 (67.21%)

* Mean ± standard deviation.

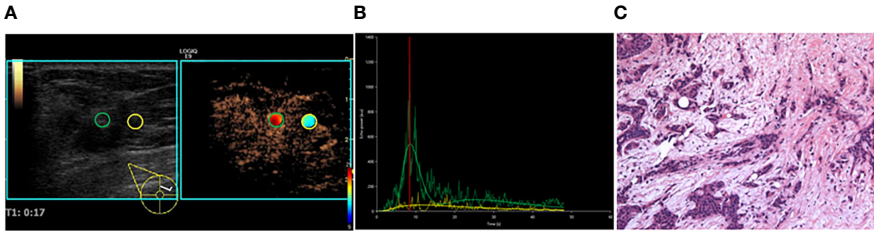


FIGURE 1
A 35-year-old female patient with invasive ductal carcinoma. **(A)** CEUS perfusion imaging using VueBox software. The ROIs were set in the breast lesion (green circle) and the surrounding breast tissue (yellow circle). The breast lesion showed iso-hyper, homogeneous enhancement, and no perfusion defect, and the radial vessel was observed. **(B)** TIC analysis showed a fast wash-in, a medium-high peak, and a rapid wash-out with the tumor. **(C)** Histopathological examination indicated a moderately differentiated (grade II), Luminal B subtyping breast cancer (HE, ×100). TIC, Time intensity curve.

enhancement features of HER-2 breast cancer were predominantly hyper-enhancement (26/31, 83.87%) and perfusion defects (27/31, 87.10%). The enhancement features of triple-negative breast cancer were predominantly clear edge (20/22, 90.91%) (Figures 2, 3). There were significant differences in the enhancement degree, enhancement edge, perfusion defects, and radial or penetrating vessels among the different molecular subtypes of breast cancer ($P < 0.05$). Radial or penetrating vessels were more common in luminal A breast cancer than in other subtypes. In addition, clear edge after enhancement was more common in the triple-negative subtype, and perfusion defect was more often found in HER-2 overexpressed breast cancer than in other subtypes ($P < 0.05$). With regard to internal homogeneity, there was no significant difference among the different molecular subtypes ($P > 0.05$) (Table 3).

Quantitative CEUS parameters of breast cancer with different pathological grades and molecular subtypes

There were significant differences in PE, mTT, TTP, WiR, WiPI, and WiWoAUC between breast cancer groups and the surrounding normal breast tissue (all $P < 0.001$). After Bonferroni correction, the PE, WiR, WiPI, and WiWoAUC for breast cancer lesions were greater, while mTT and TTP were shorter than the surrounding normal breast tissue. PE, WiR, WiPI, and WiWoAUC for the higher histological grade group were greater than the lower histological grade group (Table 4).

There were significant differences in the quantitative parameters PE, WiR, WiPI, and WiWoAUC among the different molecular subtypes of breast cancer (all $P < 0.05$). PE, WiR, WiPI, and WiWoAUC for the luminal A and luminal B subtypes were lower than triple-negative breast cancer. HER-2 overexpressed subtype had higher PE than luminal A and greater WiR than luminal A and luminal B subtypes. However, the mTT and TTP showed no statistical difference among the different molecular subtypes of breast cancer (both $P > 0.05$) (Table 5).

Diagnostic performances of different qualitative and quantitative CEUS parameters

Using the pathological results as the gold standard, ROC curves were drawn to analyze the diagnostic performance of CEUS perfusion imaging with VueBox for the pathological grades and molecular subtypes of breast cancer. The results showed that among the qualitative parameters, perfusion defect and heterogeneous enhancement were the most accurate features for higher histologic-grade breast cancer, and the areas under the ROC curve were 0.756 and 0.768, respectively. Among the quantitative parameters, PE, WiR, and WiWoAUC had the highest diagnostic performance, with 667.02, 108.55, and 6517.99 identified as the optimal cutoff values for the diagnosis of higher histologic grade breast cancer; the corresponding sensitivities were 0.895, 0.895, and 0.860; specificities were 0.673, 0.651, and 0.689; and the accuracies were 0.708, 0.689, and 0.716, respectively (Figure 4).

With a PE value of 253.96 as the threshold, hypo-enhancement, and the presence of radial or penetrating vessels used to diagnose luminal A breast cancer, the areas under the ROC curve were 0.825, 0.746, and 0.760, respectively. Using the cutoff value of 147.56 for WiPI to diagnose luminal B subtype breast cancer, the area under the ROC curve was 0.838, and the sensitivity, specificity, and accuracy were 0.925, 0.709, and 0.803. Perfusion defect and clear edge after enhancement were the best qualitative criteria for diagnosis of HER-2 overexpress and triple-negative breast cancer; the corresponding areas under the ROC curves were 0.804 and 0.905, the corresponding sensitivities were 0.871 and 0.909, specificities were 0.737 and 0.901, and the accuracies were 0.760 and 0.902, respectively. Using a WiR value of 107.81 and WiWoAUC value of 7646.07 as the cutoff values to diagnose triple-negative breast cancer, the areas under the curve were 0.932 and 0.922, the sensitivities were 0.955 and 0.955, the specificities were 0.820 and 0.857, and the accuracies were 0.836 and 0.869, respectively (Figure 5).

10.3389/fonc.2024.1320714

TABLE 2 CEUS Qualitative Features with Different Histologic grades of Breast Cancer.

Histologic grade	Enhancement Intensity			Internal Homogeneity			Enhancement edge			Perfusion defect			Radial or penetrating vessel			Enhancement scope enlarged		
	Hyper-enhancement	Iso- or hypo-enhancement		Homogeneous	Heterogeneous		clear	blurred		Present	Absent		Present	Absent		Present	Absent	
Grades I and II	47	79		83	43		24	102		26	100		45	81		77	49	
Grade III	44	13		7	50		12	45		41	16		47	10		42	15	
Total	91	92		90	93		36	147		67	116		92	91		119	64	
P Value	< 0.01			< 0.01			0.752			< 0.01			< 0.01			0.099		

Discussion

Previous studies have reported that CEUS enhancement patterns and hemodynamic changes in breast cancer can be analyzed to predict the pathological characteristics associated with breast cancer prognosis (10). However, previous studies on the preoperative assessment of breast cancer pathological grades and ER, PR, or other biomarkers' expression by CEUS mostly focused on qualitative indicators, which have certain limitations of subjectivity. In the present study, VueBox, an external perfusion analysis software, was used to comprehensively analyze and explore the value of CEUS qualitative and quantitative parameters in the preoperative assessment of the pathological grading and molecular classification of breast cancer.

The results indicated that higher histological-grade breast cancer commonly showed heterogeneous enhancement, perfusion defects, and the presence of radial or penetrating vessels. Additionally, iso- or hypo-enhancement, homogeneous enhancement, and no obvious perfusion defect were found more often in lower histological-grade breast cancer. These findings are consistent with the results of a previous report (11). Breast cancer is a vascular-dependent disease, and the differences in CEUS enhancement patterns between breast cancer lesions and the surrounding normal breast tissue are closely related to their blood perfusion and pathological characteristics. The higher the histological grade of breast cancer, the poorer the differentiation, the higher the degree of malignancy, and the more angiogenesis (12). The distribution of blood vessels in malignant tumors is uneven; there are abundant tortuous and dilated blood vessels at the edges of lesions, and immature, stenotic, and occluded new blood vessels are common, resulting in heterogeneous enhancement within the tumors (13). Perfusion defects in malignant lesions are related to rapid tumor growth and the relatively insufficient supply of oxygen and nutrients, resulting in tumor liquefaction and necrosis. On CEUS, radial or penetrating vessels may manifest as the "crab claw" sign. As we know, cells of tumors with higher histological grades could continuously secrete a large amount of vascular endothelial growth factor, which promotes the formation of new blood vessels and infiltration into surrounding normal tissue. These tumors are likely to appear as "crab claw"-like enhancement on CEUS (14).

To minimize the influence of subjective factors and individual differences on the interpretation of CEUS imaging, VueBox quantitative analysis was also used to analyze the differences in breast cancer and surrounding normal breast tissue. The results showed that the TIC curve for breast cancer lesions was characterized by rapid and hyper-perfusion. Quantitative CEUS parameters for breast cancer lesions were significantly different from those for surrounding normal breast tissue. It is probably due to the differences between the microvessel density (MVD) of the lesions and the surrounding normal glandular tissue. There is very little neovascularization in normal breast tissue, and the MVD in breast cancer lesions is significantly higher than that in normal breast tissue (15). Additionally, higher histological-grade breast cancer has more thick feeding vessels, and the neovascular wall in the tumor is incomplete. A lack of smooth muscle innervation and

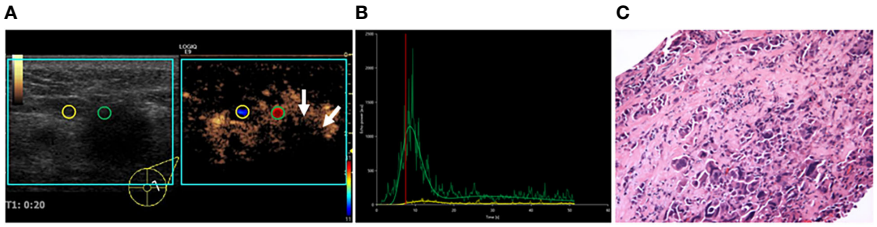


FIGURE 2
A 55-year-old female patient with invasive ductal carcinoma. **(A)** VueBox perfusion imaging showed that hyper-, heterogeneous enhancement, and perfusion defect (white arrows) were present in the breast lesion. **(B)** TIC analysis showed a fast wash-in, a higher peak, and a rapid wash-out with the tumor. **(C)** Histopathological examination indicated a poorly differentiated (grade III), Her-2 subtyping breast cancer (HE, ×100). TIC: Time intensity curve.

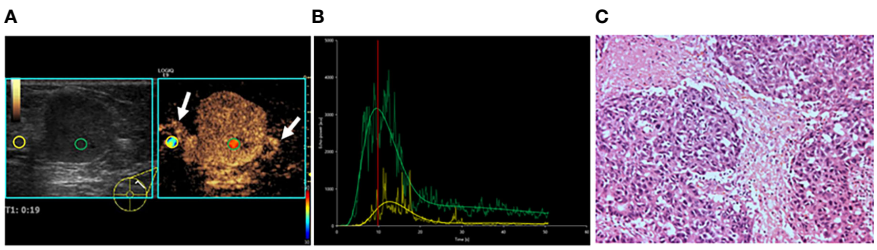


FIGURE 3
A 47-year-old female patient with invasive lobular carcinoma. **(A)** CEUS perfusion imaging showed that hyper-enhancement, clear edge, and radial vessels (white arrows) were present in the breast lesion. **(B)** TIC analysis showed a fast wash-in, a higher peak, and a rapid wash-out with a greater WiWoAUC of the tumor. **(C)** Histopathological examination indicated a poorly differentiated (grade III), triple-negative subtyping breast cancer (HE, ×100). WiWoAUC: wash-in and wash-out area under the curve.

vasomotor components and the formation of thrombi in feeding vessels in the tumor can cause a large number of microbubbles to remain in the blood vessels and eventually lead to greater PE, WiR, WiPI, and WiWoAUC and a shorter mTT and TTP values for the breast cancer lesion than for the surrounding normal glandular tissue (16). However, Li et al. (13) proposed that the longer the TTP, the smaller the WiR, and the higher the pathological grade, which is different from the present finding. The explanation for this inconsistent result may be attributable to the individual differences of patients, different ROI areas, or analysis software (17).

In this study, luminal A and luminal B subtypes of breast cancer mostly showed iso- or hypo-enhancement on CEUS. The reason for this may be the low density of microvessels, low invasiveness, and low perfusion in luminal epithelial tumors. Radial or penetrating vessels on CEUS are characteristic of luminal A breast cancer. It is speculated that luminal A breast cancer has lower Ki-67 expression, slower cell proliferation, and lower malignancy. In addition, tumor adhesion and E-cadherin expression promote the proliferation of interstitial connective tissue and inflammatory cell infiltration, leading to the formation of dense fibrosis. On CEUS, these tumors appear peripherally radial convergent, which is consistent with the burr-like appearance around masses on mammography (18).

HER-2 expression is correlated with tumor size, lymph node metastasis, and TNM stage. HER-2 overexpression often indicates poor prognosis (19). In our present study, HER-2 breast cancer mostly showed hyper-enhancement and perfusion defects. HER-2 can upregulate the expression of vascular endothelial growth factor (VEGF), increase angiogenesis, and stimulate the proliferation of microvessels around a tumor (20). The blood supply to the tumor increases, manifesting as hyper-enhancement on contrast-enhanced ultrasonography. When a tumor grows rapidly, necrosis occurs due to insufficient oxygen and nutrient supply, manifesting as perfusion defects on CEUS, which is consistent with the results reported by Liang et al. (21).

Triple-negative breast cancer has the worst prognosis and is not sensitive to endocrine therapy and targeted therapy. Previous studies have indicated that triple-negative breast cancer and benign tumors have similar appearances on conventional ultrasound (22). In this study, triple-negative breast cancer commonly showed a clear edge on CEUS, potentially relating to the compressive growth of triple-negative breast lesions, and the stromal reaction around the gland is reduced, resulting in a clear border between the tumor and the surrounding breast tissue (23).

There were different opinions of previous research regarding the correlation between quantitative CEUS parameters and the molecular expression in breast cancer. Vraka et al. (10) reported that there was no significant difference in PE, TTP, and MTT between ER-negative and ER-positive tumors. The results of another study suggested that the PE of luminal epithelial breast cancer was lower than that of HER-2 and triple-negative tumors, while the TTP of HER-2 breast cancer was shorter than that of other subtypes (21). Our results revealed that PE, WiR, WiPI, and

TABLE 3 CEUS Qualitative Features with Different Molecular Subtypes of Breast Cancer.

Molecular Subtype	Enhancement Intensity		Internal Homogeneity		Enhancement edge		Perfusion defect		Radial or penetrating vessel		Enhancement scope enlarged	
	Hyper-enhancement	Iso- or hypo-enhancement	homogeneous	Heterogeneous	clear	blurred	Present	Absent	Present	Absent	Present	Absent
Luminal A	7	43	23	27	6	44	15	35	44	6	32	18
Luminal B	39	41[a]	35	45	5	75	23	57	42	38[a]	54	26
Her-2 overexpressed	26	5[ab]	18	13	5	26	27[ab]	4	4	27[ab]	19	12
Triple-negative	19	3[ab]	14	8	20[abc]	2	2	20[c]	2	20[ab]	14	8
Total	91	92	90	93	36	147	67	116	92	91	119	64
P Value	< 0.01		0.262		< 0.01		< 0.01		< 0.01		0.930	

[a]Compared with the Luminal A subtype, $p < 0.05$.
[b]Compared with the Luminal B subtype, $p < 0.05$.
[c]Compared with the Her-2 overexpressed subtype, $p < 0.05$.

TABLE 4 Comparison of CEUS Quantitative Parameters in different histological grades of breast cancer and normal breast tissue *(Q1, Q3)*.

Group	PE (au)	mTT (s)	TTP (s)	WiR (au)	WiPI (au)	WiWoAUC (au)
Normal breast tissue	95.72 (39.19, 295.39)	39.63 (21.70, 73.87)	12.29 (8.98, 17.32)	17.67 (7.57, 51.12)	69.69 (27.06, 402.93)	1101.93 (341.01, 3869.24)
Grades I and II	840.19[a] (414.84, 2351.48)	26.56[a] (18.19, 52.72)	8.25[a] (6.57, 9.84)	151.26[a] (78.28, 269.91)	726.95[a] (277.77, 4146.50)	7400.29[a] (2672.86, 19547.71)
Grade III	2926.97[ab] (1015.94, 14848.96)	22.58[a] (14.04, 44.52)	7.74[a] (6.57, 9.84)	604.97[ab] (191.32, 3013.64)	2107.34[ab] (727.08, 9465.63)	26677.88[ab] (9280.65, 146525.00)
P Value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

[a]Compared with normal breast tissue, $p < 0.05$; [b]Compared with Grades I and II breast cancer, $p < 0.05$.
PE, peak enhancement; mTT, mean Transit time; TTP, Time to Peak; WiR, wash-in rate; WiPI, wash-in perfusion index; WiWoAUC, wash-in and wash-out areas under the curve.
au, arbitrary unit; S, second.

TABLE 5 CEUS Quantitative Parameters with Different Molecular Subtypes of Breast Cancer *(Q1, Q3)*.

Molecular Subtype	PE (au)	mTT (s)	TTP (s)	WiR (au)	WiPI (au)	WiWoAUC (au)
Luminal A	710.56 (333.10, 2584.27)	23.41 (18.04, 51.69)	8.70 (7.23, 10.04)	115.97 (62.10, 203.20)	833.74 (241.40,4106.77)	7617.21 (2104.91,38584.92)
Luminal B	876.70 (442.76, 3357.51)	23.56 (16.87, 44.9175)	7.975 (6.98, 10.3625)	175.97 (83.92, 427.59) [a]	817.85 (325.13,4424.32)	10368.69 (3927.83,45378.18)
Her-2 overexpressed	2082.53 (729.85, 6682.69)[a]	29.00 (15.11, 54.04)	8.36 (7.28, 10.72)	629.8 (197.28,1361.85)[ab]	1762.22 (465.52,22005.59)	12184.65 (4372.72,36945.39)
Triple-negative	2946.19 (1565.76, 16195.12)[ab]	29.495 (14.43, 52.09)	7.74 (6.40, 9.62)	879.26 (230.51,5049.16)[ab]	2116.12 (1132.53,13369.86)[a,b]	34138.295 (11401.22,168665.78)[a,b]
P Value	< 0.001	0.842	0.478	< 0.001	< 0.001	0.007

[a]Compared with Luminal A subtype, $p < 0.05$; [b]Compared with Luminal B subtype, $p < 0.05$.
PE, peak enhancement; mTT, mean Transit time; TTP, Time to Peak; WiR, wash-in rate; WiPI, wash-in perfusion index; WiWoAUC, wash-in and wash-out areas under the curve.
au, arbitrary unit; S, second.

WiWoAUC for the luminal A and luminal B subtypes were lower than triple-negative breast cancer. HER-2 overexpressed subtype had higher PE than luminal A and greater WiR than luminal A and luminal B subtypes. Our results concur with those of Wen B et al. (24) in that HER-2 overexpressed and triple-negative breast lesions can secrete more vascular endothelial growth factor, leading to higher angiogenesis and vascular permeability and significant hyper-perfusion situation in the tumors.

This study had certain limitations. First, this was a retrospective study, and there may be some selection bias. Second, the selected ROIs were lesion areas with the strongest enhanced area, thus not fully representing blood perfusion in entire lesions. Third,

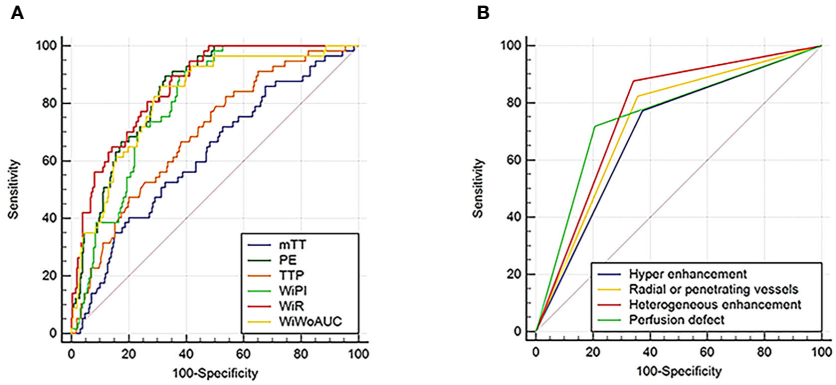


FIGURE 4
ROC curves for CEUS quantitative **(A)** and qualitative **(B)** parameters in the diagnosis of higher histologic grade (grade III) breast cancer. ROC, Receiver operating characteristic; CEUS, Contrast enhanced ultrasound.

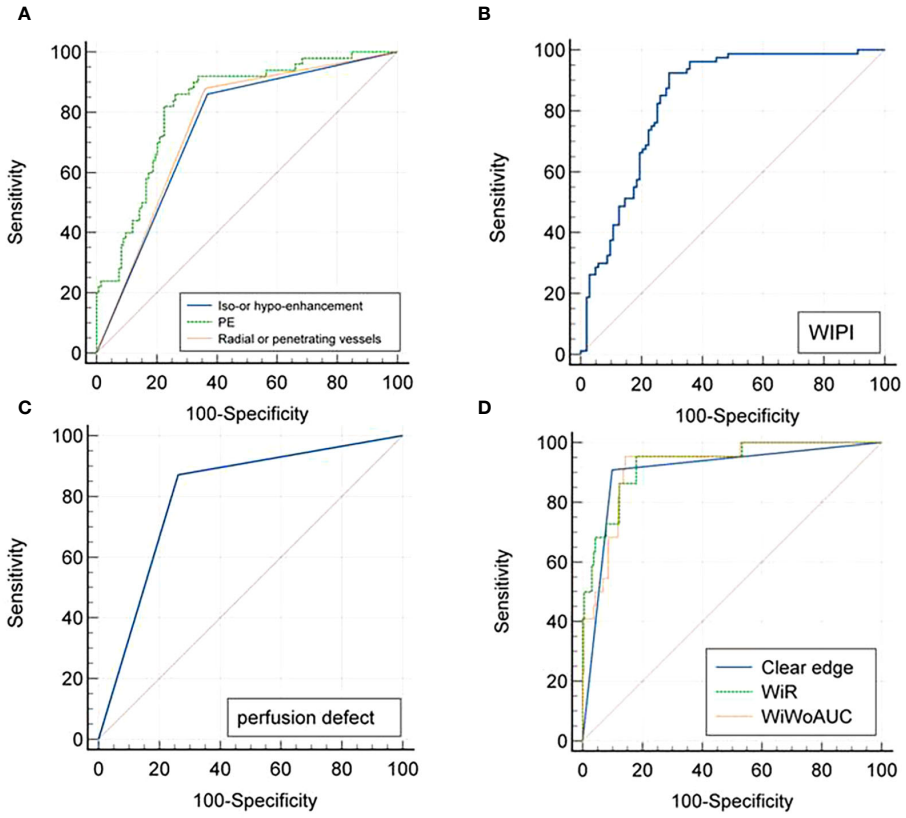


FIGURE 5
ROC curves for CEUS diagnoses of **(A)** Luminal A, **(B)** Luminal B, **(C)** Her-2 overexpressed, and **(D)** Triple-negative subtyping breast cancer. ROC, Receiver operating characteristic; CEUS, Contrast-enhanced ultrasound.

univariate analysis was used in the present study and the sample size was relatively small; the combined value of CEUS quantitative and qualitative parameters in predicting the pathological characteristics of breast cancer needs to be calculated in a multivariate regression analysis and verified in future multicenter and large-scale studies.

In conclusion, there were differences in the qualitative features and quantitative parameters of CEUS for breast cancer with different pathological grades and molecular subtypes. Contrast-enhanced ultrasound may be used to non-invasively predict the histological characteristics of breast cancer.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by The Ethics Committee of the People's Hospital of Guangxi Zhuang Autonomous Region. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

L-LL: Data curation, Formal analysis, Writing – original draft, Validation. Q-LS: Writing – original draft, Data curation, Formal analysis, Validation. Y-XD: Data curation, Writing – original draft, Validation. W-WG: Methodology, Validation, Writing – original draft. H-ML: Validation, Formal analysis, Writing – review & editing. QH: Funding acquisition, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Guangxi Key Research and Development project and the National Natural Science Foundation of China (No. 82160339, 81660292). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin*. (2021) 71:209–49. doi: 10.3322/caac.21660

2. Tsang JYS, Tse GM. Molecular classification of breast cancer. *Adv Anat Pathol*. (2020) 27:27–35. doi: 10.1097/PAP.0000000000000232

3. Rakha EA, Green AR. Molecular classification of breast cancer: what the pathologist needs to know. *Pathology*. (2017) 49:111–9. doi: 10.1016/j.pathol.2016.10.012

4. Wiesinger I, Wiggermann P, Zausig N, Beyer LP, Salzberger B, Stroszczynski C, et al. Percutaneous treatment of Malignant liver lesions: evaluation of success using contrast- enhanced ultrasound (CEUS) and perfusion software. *Ultraschall der Med*. (2018) 39:440–7. doi: 10.1055/s-0043-119353

5. Wiesinger I, Jung F, Jung EM. Contrast-enhanced ultrasound (CEUS) and perfusion imaging using VueBox®. *Clin Hemorheol Microcirc*. (2021) 78:29–40. doi: 10.3233/CH-201040

6. Dong Y, Koch JBH, Löwe AL, Christen M, Wang WP, Jung EM, et al. VueBox® for quantitative analysis of contrast-enhanced ultrasound in liver tumors. *Clin Hemorheol Microcirc*. (2022) 80:473–86. doi: 10.3233/CH-211261

7. Jung EM, Jung F, Stroszczynski C, Wiesinger I. Quantification of dynamic contrast-enhanced ultrasound (CEUS) in non-cystic breast lesions using external perfusion software. *Sci Rep*. (2021) 11:17677. doi: 10.1038/s41598-021-96137-6

8. Tavassoli FA, Devilee P. Pathology and Genetics Tumors of the Breast and Female Genital Organs. In: *World Health Organization classification of tumors*. IARC Press, Lyon (2003). p. 19–23.

9. Khaled H, Gamal H, Lotayef M, Knauer M, Thürliman B. The st. Gallen international expert consensus conference on the primary therapy of early breast cancer 2017: Egyptian view. *Breast Cancer Res Treat*. (2018) 172:545–50. doi: 10.1007/s10549-018-4945-1

10. Vraka I, Panourgias E, Sifakis E, Koureas A, Galanis P, Dellaportas D, et al. Correlation between contrast-enhanced ultrasound characteristics (Qualitative and quantitative) and pathological prognostic factors in breast cancer. *In Vivo*. (2018) 32:945–54. doi: 10.21873/invivo.11333

11. Cao XL, Bao W, Zhu SG, Wang LH, Sun MH, Wang L, et al. Contrast-enhanced ultrasound characteristics of breast cancer: correlation with prognostic factors. *Ultrasound Med Biol*. (2014) 40:11–7. doi: 10.1016/j.ultrasmedbio.2013.08.014

12. Karimzadeh P, Faghih Z, Rahmani N, Eghbali F, Razmkhah M. Quantification of angiogenic factors and their clinicopathological associations in breast cancer. *Eur Cytokine Netw*. (2020) 31:68–75. doi: 10.1684/ecn.2020.0447

13. Li X, Li Y, Zhu Y, Fu L, Liu P. Association between enhancement patterns and parameters of contrast-enhanced ultrasound and microvessel distribution in breast cancer. *Oncol Lett*. (2018) 15:5643–9. doi: 10.3892/ol.2018.8078

14. Zhao YX, Liu S, Hu YB, Ge YY, Lv DM. Diagnostic and prognostic values of contrast-enhanced ultrasound in breast cancer: a retrospective study. *Onco Targets Ther*. (2017) 10:1123–9. doi: 10.2147/OTT.S124134

15. Ribatti D, Nico B, Ruggieri S, Tamma R, Simone G, Mangia A. Angiogenesis and antiangiogenesis in triple-negative breast cancer. *Transl Oncol*. (2016) 9:453–7. doi: 10.1016/j.tranon.2016.07.002

16. Li W, Zhou Q, Xia S, Wu Y, Fei X, Wang Y, et al. Application of contrast-enhanced ultrasound in the diagnosis of ductal carcinoma in situ: analysis of 127 cases. *J Ultrasound Med*. (2020) 39:39–50. doi: 10.1002/jum.15069

17. Li J, Yuan M, Yang L, Guo L. Correlation of contrast-enhanced ultrasound features with prognostic factors in invasive ductal carcinomas of the breast. *Jpn J Radiol*. (2020) 38:960–7. doi: 10.1007/s11604-020-00994-6

18. Erber R, Hartmann A. Histology of luminal breast cancer. *Breast Care (Basel)*. (2020) 15:327–36. doi: 10.1159/000509025

19. Shi P, Chen C, Yao Y. Correlation between HER-2 gene amplification or protein expression and clinical pathological features of breast cancer. *Cancer Biother Radiopharm*. (2019) 34:42–6. doi: 10.1089/cbr.2018.2576

20. Hasan R, Bhatt D, Khan S, Khan V, Verma AK, Anees A, et al. Association of her-2 expression and clinicopathological parameters in colorectal carcinoma in Indian population. *Open Access Maced J Med Sci*. (2018) 7:6–11. doi: 10.3889/oamjms.2019.008

21. Liang X, Li Z, Zhang L, Wang D, Tian J. Application of contrast-enhanced ultrasound in the differential diagnosis of different molecular subtypes of breast cancer. *Ultrason Imaging*. (2020) 42:261–70. doi: 10.1177/0161734620959780

22. Li JW, Cao YC, Zhao ZJ, Shi ZT, Duan XQ, Chang C, et al. Prediction for pathological and immunohistochemical characteristics of triple-negative invasive

breast carcinomas: the performance comparison between quantitative and qualitative sonographic feature analysis. *Eur Radiol*. (2022) 32:1590–600. doi: 10.1007/s00330-021-08224-x

23. Wang D, Zhu K, Tian J, Li Z, Du G, Guo Q, et al. Clinicopathological and ultrasonic features of triple-negative breast cancers: A comparison with hormone receptor-positive/human epidermal growth factor receptor-2-negative breast cancers. *Ultrasound Med Biol*. (2018) 44:1124–32. doi: 10.1016/j.ultrasmedbio.2018.01.013

24. Wen B, Kong W, Zhang Y, Xue H, Wu M, Wang F. Association between contrast-enhanced ultrasound characteristics and molecular subtypes of breast cancer. *J Ultrasound Med*. (2022) 41:2019–31. doi: 10.1002/jum.15886

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 22 May 2024
DOI 10.3389/fonc.2024.1356173

OPEN ACCESS

EDITED BY
Sharon R. Pine,
University of Colorado Anschutz Medical Campus, United States

REVIEWED BY
Savannah Partridge,
University of Washington, United States
Nathalie Just,
Copenhagen University Hospital, Denmark

*CORRESPONDENCE
David L. Buckley
D.L.Buckley@leeds.ac.uk

RECEIVED 15 December 2023
ACCEPTED 07 May 2024
PUBLISHED 22 May 2024

CITATION
Almutlaq ZM, Bacon SE, Wilson DJ, Sharma N, Dondo T and Buckley DL (2024) The relationship between parameters measured using intravoxel incoherent motion and dynamic contrast-enhanced MRI in patients with breast cancer undergoing neoadjuvant chemotherapy: a longitudinal cohort study. *Front. Oncol.* 14:1356173.
doi: 10.3389/fonc.2024.1356173

The relationship between parameters measured using intravoxel incoherent motion and dynamic contrast-enhanced MRI in patients with breast cancer undergoing neoadjuvant chemotherapy: a longitudinal cohort study

Zyad M. Almutlaq[1,2], Sarah E. Bacon[3], Daniel J. Wilson[3], Nisha Sharma[4], Tatendashe Dondo[5] and David L. Buckley[1*]

[1]Biomedical Imaging, Leeds Institute of Cardiovascular and Metabolic Medicine (LICAMM), University of Leeds, Leeds, United Kingdom, [2]Radiological Sciences Department, College of Applied Medical Sciences, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia, [3]Department of Medical Physics & Engineering, Leeds Teaching Hospitals National Health Service (NHS) Trust, Leeds, United Kingdom, [4]Department of Radiology, Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom, [5]Clinical and Population Sciences Department, LICAMM, University of Leeds, Leeds, United Kingdom

Purpose: The primary aim of this study was to explore whether intravoxel incoherent motion (IVIM) can offer a contrast-agent-free alternative to dynamic contrast-enhanced (DCE)-MRI for measuring breast tumor perfusion. The secondary aim was to investigate the relationship between tissue diffusion measures from DWI and DCE-MRI measures of the tissue interstitial and extracellular volume fractions.

Materials and methods: A total of 108 paired DWI and DCE-MRI scans were acquired at 1.5 T from 40 patients with primary breast cancer (median age: 44.5 years) before and during neoadjuvant chemotherapy (NACT). DWI parameters included apparent diffusion coefficient (ADC), tissue diffusion (D_t), pseudo-diffusion coefficient (D_p), perfused fraction (f), and the product $f{\times}D_p$ (microvascular blood flow). DCE-MRI parameters included blood flow (F_b), blood volume fraction (v_b), interstitial volume fraction (v_e) and extracellular volume fraction (v_d). All were extracted from three tumor regions of interest (whole-tumor, ADC cold-spot, and DCE-MRI hot-spot) at three MRI visits: pre-treatment, after one, and three cycles of NACT. Spearman's rank correlation was used for assessing between-subject correlations (r), while repeated measures correlation was employed to assess within-subject correlations (r_{rm}) across visits between DWI and DCE-MRI parameters in each region.

Results: No statistically significant between-subject or within-subject correlation was found between the perfusion parameters estimated by IVIM and DCE-MRI

(f versus v_b and fxD_p versus F_b; P=0.07–0.81). Significant moderate positive between-subject and within-subject correlations were observed between ADC and v_e (r=0.461, r_{rm}=0.597) and between D_t and v_e (r=0.405, r_{rm}=0.514) as well as moderate positive within-subject correlations between ADC and v_d and between D_t and v_d (r_{rm}=0.619 and 0.564, respectively) in the whole-tumor region.

Conclusion: No correlations were observed between the perfusion parameters estimated by IVIM and DCE-MRI. This may be attributed to imprecise estimates of fxD_p and v_b, or an underlying difference in what IVIM and DCE-MRI measure. Care should be taken when interpreting the IVIM parameters (f and $f\times D_p$) as surrogates for those measured using DCE-MRI. However, the moderate positive correlations found between ADC and D_t and the DCE-MRI parameters v_e and v_d confirms the expectation that as the interstitial and extracellular volume fractions increase, water diffusion increases.

KEYWORDS

breast cancer, intravoxel incoherent motion, dynamic contrast enhanced MRI, perfusion, repeated measures, correlations

1 Introduction

Breast cancer is one of the most prevalent cancers affecting women globally, with about 2.3 million women diagnosed with the disease and 685,000 deaths in 2020 (1). Patients with primary breast cancer are often treated with neoadjuvant chemotherapy (NACT) to downsize the tumor and increase the probability of breast-conserving surgery (2). A non-invasive imaging technique that can provide information on tumor cellularity and perfusion during treatment would be beneficial, as reduced cellularity and perfusion are promising indicators of patient response to treatment (3).

Patients with breast cancer undergoing NACT often undergo repeated dynamic contrast-enhanced (DCE) MRI scans for treatment monitoring (4). DCE-MRI is a widespread technique that can provide information on tumor perfusion and cellularity through serial MRI scans acquired before and after the injection of a gadolinium-based contrast agent (3). Furthermore, quantitative estimation of perfusion-related parameters of breast tumors, including tumor blood flow (F_b), blood volume fraction (v_b), along with hemodynamic and cellularity-related parameters: capillary permeability–surface area product (PS); interstitial volume fraction (v_e), and extracellular volume fraction (v_d; calculated from the combination of blood volume and interstitial volume fractions) can be achieved by employing a recently developed DCE-MRI technique (5). However, certain safety concerns exist regarding gadolinium administration, particularly in patients with cancer who undergo repeated contrast-enhanced scans (6). Therefore, alternative imaging techniques that can provide equivalent perfusion and cellularity-related measurements without administering a contrast agent are of interest.

Conventional diffusion-weighted imaging (DWI), which is not used generally in breast cancer imaging, can be employed in oncology treatment response monitoring through the apparent diffusion coefficient (ADC). The ADC measures the diffusivity of water molecules in the tissue and is assumed to serve as an indicator of cellular density. As such, as tumor cellularity decreases in response to treatment, the ADC value increases (7). The ADC is therefore expected to be directly proportional to the DCE-MRI measurements of the tissue's interstitial and extracellular volume fractions. However, few studies have examine this relationship, and one study in breast tumors has challenged the expectation suggesting that the ADC is incompletely understood (8, 9). Also, blood in the microcirculation can contaminate the DWI signal decay, contributing to the ADC value (10). A technique that can potentially address the problems affecting both DCE-MRI and DWI is intravoxel incoherent motion (IVIM), an advanced form of DWI. It has been proposed that IVIM enables simultaneous assessment of tissue diffusion and perfusion by separating the effects of the microcirculation of blood in the capillary network (so-called pseudo-diffusion) from water diffusion in the rest of the tissue. This method requires DWI acquisitions with multiple b-values (low and high) and fits a bi-exponential model to the data to estimate the diffusion-related parameter D_t (tissue diffusion) and perfusion-related parameters, including D_p (the pseudo-diffusion coefficient), f (the perfused fraction), and their product $f\times D_p$ (microvascular blood flow) (10, 11).

In the past decade, growing interest in exploring the potential applications of IVIM in breast tumors has produced studies differentiating benign and malignant tumors (12, 13). IVIM perfusion-related parameters have also shown some promise for evaluating breast tumor response to NACT (14–16). This in turn

has reopened the question of whether IVIM could be used as a contrast-agent-free alternative to DCE-MRI for measuring breast tumor perfusion. Few studies have investigated the correlations between IVIM and DCE-MRI perfusion-related parameters in breast tumors and have produced contradictory results (17–19). These studies examined correlations at a single visit; however, a correlation between perfusion parameter changes caused by treatment is meaningful and suggests that IVIM could be a contrast-agent-free surrogate to the DCE-MRI method in monitoring serial changes in tumor perfusion. Further, none of these studies provided an absolute estimation of tumor blood flow; they did not perform a direct comparison with the IVIM parameter purported to measure microvascular blood flow ($f \times D_p$).

The primary aim of this study was to investigate whether IVIM and DCE-MRI perfusion-related parameters correlate and whether IVIM can offer a contrast-agent-free alternative to DCE-MRI for monitoring serial changes in tumor perfusion. The DCE-MRI data were analyzed to estimate absolute tumor blood flow, blood volume fraction, capillary permeability–surface area product, interstitial volume fraction, and extracellular volume fraction (5). This study assesses both between-subject and within-subject repeated measures correlations between the perfusion parameters estimated by IVIM and DCE-MRI (specifically perfusion fraction versus blood volume fraction and microvascular blood flow versus blood flow) in a cohort of patients with breast cancer imaged before treatment and after one and three cycles of NACT. Analyzing both correlations is valuable; between-subject correlation reveals the potential for estimating DCE-MRI perfusion parameters using IVIM at a given time, whereas within-subject repeated measures correlations indicate the potential for estimating change in DCE-MRI perfusion parameters using IVIM when assessing longitudinal changes in the same patient. The secondary aim of this study was to examine the correlation between tissue diffusion measures from DWI and DCE-MRI measures of the tissue's interstitial and extracellular volume fractions. This would improve the understanding of tissue diffusion measures and their changes in response to treatment further, which are of interest for translation into breast cancer imaging as markers of treatment response (20).

2 Materials and methods

2.1 Patients

The prospective study had local research ethics committee approval, and written informed consent was obtained from all subjects. The eligibility criteria for patient inclusion were: 1) 18 years of age and older; 2) pathological confirmation of primary invasive breast cancer through a core needle biopsy; and 3) scheduled to undergo NACT. Patients who had impaired kidney function or contraindications to MRI were considered ineligible. Recruited patients underwent a standardized NACT regimen consisting of three cycles of epirubicin (90 mg/m^2) and cyclophosphamide (600 mg/m^2) (one cycle every three weeks), followed by three cycles of docetaxel (100 mg/m^2, one cycle every three weeks). Patients with tumors positive for human epidermal growth factor receptor 2 were treated with trastuzumab and/or pertuzumab alongside docetaxel.

2.2 Image acquisition

MRI scans were performed at baseline (pre-treatment) and after one and three (mid-treatment) cycles of NACT. All images were acquired using a 1.5-T MRI scanner (Aera; Siemens) with the patient in a head-first prone position. A dedicated 16-channel breast coil (Sentinelle; Siemens) was used to image the breasts, and a flexible array coil, placed on the patient's back, was employed to increase the signal from the descending aorta (21). The scanning protocol included axial T_2-weighted turbo spin-echo, axial T_1-weighted 3D spoiled gradient echo, inversion recovery, DWI, and DCE-MRI sequences.

The axial DWI was acquired using a spectral attenuated inversion recovery fat-suppressed, 2D single-shot spin-echo echo-planar imaging sequence (repetition time/echo time: 7200/59 ms, flip angle: 90°, field of view: 340×136×169 mm, matrix size: 280×116×34, slice thickness: 4 mm, acceleration factor: 2, acquisition time: 5 min 31 s) performed at six b-values (0, 50, 100, 200, 400, and 800 s/mm^2; gradient system: strength 45 mT/m, slew rate 200 T/m/s)). The high b-value of 800 s/mm^2 was chosen in line with consensus recommendations for breast DWI (22). ADC maps were generated by the scanner software after DWI acquisition. This step was followed by a 3D non-selective inversion recovery -prepared spoiled gradient echo sequence (repetition time/echo time: 2.8/0.93 ms, flip angle: 8°, field of view: 340×340×180 mm, matrix size: 128×128×36, slice thickness: 5 mm, acceleration factor: 2, inversion recovery - repetition time: 3000 ms, overall acquisition time: 4 min 20 s), performed at four inversion times (100, 600, 1200 and 2800 ms) to estimate T_1. Both breasts, the aortic arch and part of the descending aorta were included in the field of view (21).

Afterwards, interleaved high temporal resolution (HTR) and high spatial resolution (HSR) DCE-MRI sequences were employed (5). The dynamic series consisted of 93 HTR images interleaved with 8 HSR images acquired as follows (10×HTR, 1×HSR, 43×HTR, [1×HSR, 5×HTR] repeated seven times, and finally 5×HTR). The HTR dynamic images were acquired using a T_1-weighted 3D spoiled gradient echo sequence (repetition time/echo time: 2.37/0.73 ms, flip angle: 25°, field of view: 340×340 ×180 mm, matrix size: 128×128×36, slice thickness: 5 mm, acceleration factor: 2×2, acquisition time: 2 s). For the HSR images, a fat-suppressed T_1-weighted 3D spoiled gradient echo sequence (repetition time/echo time: 4.1/1.2 ms, flip angle: 10°, field of view: 340×340 ×180 mm, matrix size: 384×384×128, slice thickness: 1.4 mm, acceleration factor: 3, and acquisition time: 36 s) was employed. The HTR and HSR images were acquired with the same geometry as the inversion recovery sequence. Using an automated power injector (Spectris Solaris EP), gadolinium-based contrast agent (Dotarem, Guerbet Laboratories) was administered intravenously (0.1 mmol/kg) at the

start of the eleventh HTR-DCE-MRI image, followed by saline (20 ml at a rate of 3 ml/s). A second inversion recovery T1 estimate (bookend) was performed (23) after all eight HSR (and 88 HTR images) images were obtained. Then, the last five HTR images were acquired.

2.3 Image analysis

MRI data were processed using in-house programs developed in MATLAB (MathWorks, USA). The DWI images (including ADC maps) were rigidly aligned to the corresponding HSR, HTR and inversion recovery images to match the slice position with no interpolation of the DWI data, and HTR and HSR subtraction images were generated to improve tumor visibility. The location of the largest tumor for each patient was determined using HSR DCE-MRI images from the baseline MRI, confirmed by a breast radiologist. Then, a whole-tumor region of interest was generated using a 3D region-growing algorithm based on the enhanced tumor's signal intensity in HSR subtraction images, while avoiding obvious necrotic areas manually. Two smaller single-slice regions of interest (5×5 pixels) within the whole-tumor region were generated to reduce the possibility of tumor heterogeneity compromising subsequent correlation analysis. These small regions comprised the region with the lowest ADC on the ADC map (cold-spot region) (22) and the region with the highest SI on the HTR subtraction images (hot-spot region). All three regions were propagated to the corresponding DWI, inversion recovery and HTR images for further analysis (Figure 1). The spatial location of the smaller regions generated for each tumor were allowed to vary at each MRI visit as the tumor responded to NACT.

For DCE-MRI, these three regions were used to estimate T1 relaxation-times from both sets of inversion recovery images. A further region of interest was drawn in the descending aorta to generate signal intensity -time curves and estimate T1 before and after gadolinium-based contrast agent injection for measurement of the arterial input function (21). The signal intensity -time data were converted to gadolinium-based contrast concentration-time using a bookend T1 correction with an iterative scheme (21, 23). A two–compartment exchange model was fitted to the DCE-MRI data, and tumor blood flow, blood volume fraction, capillary permeability–surface area product and interstitial volume fraction were estimated (24). Then, the extracellular volume fraction (the sum of interstitial and blood volume fractions) was calculated. For each region of

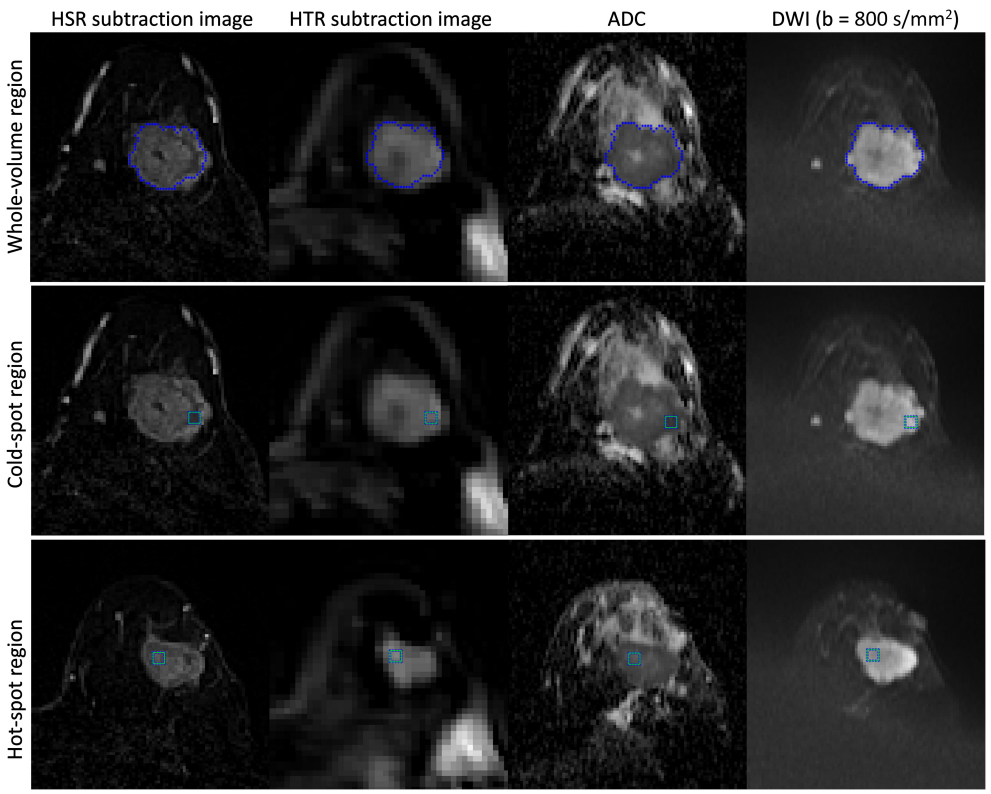


FIGURE 1
Example of seeding a tumor in a 45-year-old woman with invasive ductal carcinoma in the left breast and generating three regions of interest (whole-tumor, cold-spot, and hot-spot). First, the tumor was seeded on the HSR subtraction images, and the whole-tumor region of interest was generated (top row). Then, two smaller regions of interest (5×5 pixels) were generated within the whole-tumor region (cold-spot (middle row) and hot-spot regions (bottom row)). The whole-tumor region of interest encompasses all the slices in which the tumor appears, while the cold-spot and hot-spot regions originate in only a single slice (not necessarily the same slice). All three regions were propagated to the corresponding DWI, inversion recovery, and HTR images for further analysis.

interest, a tissue uptake model (described by parameters blood flow, blood volume fraction, and capillary permeability–surface area product) and a one–compartment model (described by parameters blood flow and extracellular volume fraction) were also fitted to the DCE-MRI data (24). The final model to use in the correlation analysis was selected based on the corrected Akaike information criterion test (cAIC) to evaluate which model best fits the data (25, 26).

For DWI, the mean signal intensity for each b–value was extracted from the three regions (27). The IVIM parameters were estimated by fitting the bi-exponential model to the mean signal intensity vs b-value data using an over-segmented approach, where tissue diffusion and perfusion fraction were estimated first and then pseudo-diffusion (28). The monoexponential model was also fitted to the mean signal intensity vs b-value, and the ADC value for each region of interest was estimated (28). The two model equations are detailed in the Supplementary Material (Appendix A). This step was conducted blinded to the DCE-MRI parameter values. Further, a simulation study was performed to assess the bias and precision of the IVIM parameter estimates with 6 b-values in comparison with 12 b-values (methods and results are provided in Appendix B, Supplementary Material).

2.4 Statistical analysis

Due to the non-normal data distribution, the DWI and DCE-MRI data were summarized using the median (interquartile range). Friedman's test with Bonferroni correction (Bonferroni *post hoc* test) was performed for each parameter from the baseline MRI to determine whether parameter differences existed between the three regions of interest (whole-tumor, cold-spot, and hot-spot). To determine the between-subject correlation between IVIM and DCE-MRI parameters for each region, the mean value of each parameter for each patient was calculated by dividing the sum of parameter values from all MRI visits by the number of times the parameter was estimated; then, the parameter value for each visit where the parameter was estimated was replaced by its subject mean. The weighted correlation coefficient, r, was calculated between the mean DWI and DCE-MRI parameters for each region of interest using the Spearman's rank correlation test (29) ($r<0.2$, very weak; $0.2\leq r<0.4$, weak; $0.4\leq r<0.7$, moderate; $0.7\leq r<0.9$, strong; $r\geq0.9$, very strong correlation) (30). This statistical method was followed to exploit the properties of data with multiple measures while addressing the issue of non-independence among observations and the impact of NACT (29). Statistical analyses were performed using SPSS software for Windows (v.25.0, Chicago, IL). All tests were two-sided, and a p-value of less than 0.05 was considered statistically significant.

To determine the correlation between changes in the IVIM and DCE-MRI parameters induced by treatment, the repeated measures correlation test (rmcorr) was utilized via the rmcorr-shiny app (31, 32). The rmcorr-shiny app computes a repeated measures correlation coefficient (r_{rm}) that considers the dependence between repeated measurements. This analysis involves determining the correlation between two parameters while accounting for between-subject variation. The rmcorr-shiny app fits separate parallel lines to each patient's data utilizing a shared slope but permitting the intercept to differ per patient. The orientation of these parallel lines represents the correlation's sign (positive or negative), while the slope denotes the correlation's magnitude.

The results of repeated measures correlation for each region were summarized in tables as: $r_{rm,}$ degrees of freedom, 95% confidence interval, and a p-value. The 95% confidence interval for each r_{rm} were determined using bootstrapping with 1000 resamples. The degrees of freedom (df) were computed based on the formula $df = N(k-1) - 1$, where N is the total number of patients and k is the (average) number of repeated measures per patient (31). The rmcorr test was initially conducted to identify statistically significant results (P-value < 0.05), then bootstrapped 95% confidence intervals were calculated. A correlation result was considered meaningful and significant only if the magnitude of the correlation coefficient was ≥ 0.4, the P-value was less than 0.05, and the bootstrapped 95% confidence intervals excluded zero. Since this is a preliminary exploration study focusing on hypothesis generation, P-values for the correlation tests were reported as raw values and were not corrected for multiple comparisons. An upper estimate of the repeatability of the DWI and DCE-MRI parameters was calculated from a subset of baseline and cycle 1 studies (details included in Appendix C, Supplementary Material).

3 Results

In this study, 40 female patients were eligible and enrolled between August 2015 and April 2018 (median age 44.5 (39, 53) years). MRI data were obtained for all patients at baseline and 37 patients after one and three cycles of NACT (three withdrew following baseline MRI). However, the MRI data acquired after three NACT cycles from two patients were excluded from the analysis because no tumor was apparent on their MRI scans. This exclusion resulted in 112 MRI studies with DWI and DCE-MRI acquisitions. Table 1 presents the clinical characteristics of the patients.

Four DCE-MRI scans—two at baseline and two after three NACT cycles—were excluded because one patient could not tolerate the whole imaging protocol, two had technical issues (the back coil was switched on and off sporadically), and one moved during the DCE-MRI acquisition, leaving 108 studies with paired DWI and DCE-MRI data acquisitions (Figure 2). Based on the cAIC results, 75 DCE-MRI data sets were analyzed using the two–compartment exchange model, 20 using the tissue uptake model, and 13 using the one–compartment model.

Smaller regions of interest (cold-spot and hot-spot regions) were generated from 91 out of 108 studies: 34 at the baseline, 33 after one NACT cycle, and 24 after three NACT cycles (some tumors shrank below the 5x5 pixels threshold during NACT). DCE

10.3389/fonc.2024.1356173

TABLE 1 Clinical characteristics of all the enrolled patients.

Characteristic	Number or Median (Interquartile range)
Number of patients	40
Age (years)	44.5 (38.8, 53.0)
Tumor volume (cm^3)	
At baseline (N= 40)	5.45 (2.16, 16.27)
After one cycle of NACT (N= 37)	4.1 (1.57, 8.83)
After three cycles of NACT (N= 35)	2.15 (0.53, 5.8)
Tumor grade	
II	15
III	25
Tumor type	
Invasive ductal carcinoma	38
Inflammatory breast cancer	1
Mucinous carcinoma	1
Estrogen receptor status	
Positive (+)	28
Negative (-)	12
Progesterone receptor status	
Positive (+)	18
Negative (-)	20
Not evaluable	2
Human epidermal growth factor 2 status	
Positive (+)	15
Negative (-)	25

data sets were fitted using the two-compartment exchange/tissue-uptake/one-compartment models for 68/13/10 cold-spot regions and for 72/9/10 hot-spot regions.

For DWI data analysis, there were a number of cases where estimates of the IVIM parameters pseudo-diffusion and perfusion fraction reached one of their limiting values, and these parameters were excluded from the statistical analyses (2 cases from the whole-tumor region, 8 from the cold-spot region, and 4 from the hot-spot region).

3.1 Estimated DWI and DCE-MRI parameters from the three regions at baseline

There were significant differences between the parameter values estimated in whole-tumor, cold-spot, and hot-spot regions for all DWI and DCE-MRI parameters, with the exception of pseudo-diffusion and extracellular volume fraction (P=0.88 and 0.2, respectively). Detailed results, including pairwise comparisons (Bonferroni-corrected), are presented in the Supplementary Material (Supplementary Table A1).

3.2 Correlation between averaged DWI and DCE-MRI parameters from three MRI visits (between-subject correlation)

No significant correlations were discovered between the IVIM and DCE-MRI perfusion-related parameters (perfusion fraction with blood volume fraction, and microvascular blood flow with blood flow) in the three tumor regions (P=0.146–0.379, Table 2, Supplementary Table A2, Supplementary Table A3). However, for whole-tumor regions, ADC exhibited a significant moderate positive correlation with tumor T_1 and interstitial volume fraction (r = 0.603 and 0.461, respectively). Similarly, D_t demonstrated a significant moderate positive correlation with tumor T_1 and interstitial volume fraction (r = 0.631 and 0.405, respectively). (Figure 3, Table 2).

In the cold-spot regions, significant moderate positive correlations were found between tumor T_1 and both measures of tissue diffusion ADC and D_t (r= 0.632 and 0.588, respectively). pseudo-diffusion demonstrated a significant moderate negative correlation with blood flow (r = -0.400, Supplementary Table A2). In hot-spot regions, ADC and D_t displayed significant moderate positive correlations with tumor T_1 (r=0.520 and 0.460, respectively, Supplementary Table A3).

3.3 Repeated measures correlations between DWI and DCE-MRI parameters (within-subject correlation)

Table 3 lists the repeated measures correlation results computed between the DWI and DCE-MRI parameters estimated from the whole-tumor regions of interest. No statistically significant correlations were discovered between the IVIM and DCE-MRI perfusion-related parameters of the study's primary interest (perfusion fraction versus blood volume fraction and microvascular blood flow versus blood flow; P=0.815 and 0.229, respectively). However, ADC and D_t displayed significant moderate positive correlations with interstitial volume fraction (r_{rm}=0.597 and 0.514, respectively) and extracellular volume fraction (r_{rm}=0.619 and 0.564, respectively) (Figure 3).

The median DWI and DCE-MRI parameter values estimated at the three MRI visits from the cold-spot and hot-spot regions exhibited patterns similar to those of the whole-tumor regions but with much more variability (Figure 4, Supplementary Figure A1); repeated measures correlation results in the cold-spot and hot-spot regions are presented in Supplementary Data only (Supplementary Table A4, Supplementary Table A5).

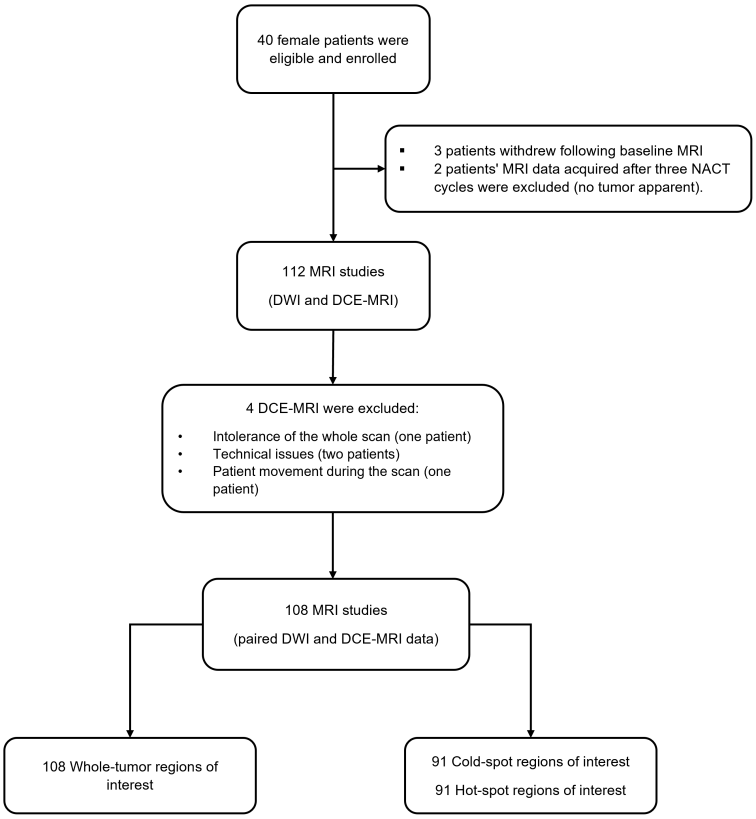


FIGURE 2
The flow chart illustrates the number of recruited patients, excluded DCE-MRI data, and the final number of MRI studies with paired DWI and DCE-MRI data acquisitions.

4 Discussion

Despite the examination of 108 paired DWI and DCE-MRI datasets, no statistically significant between-subject or within-subject repeated measures correlations were found between the IVIM and DCE-MRI perfusion parameters of the study's primary interest (perfusion fraction versus blood volume fraction and microvascular blood flow versus blood flow). These findings align with previous breast cancer studies, which also found no correlation between any IVIM perfusion parameters and DCE-MRI parameter related to perfusion, K^{trans} (transfer constant) (18, 19). K^{trans} may not solely reflect tumor blood flow but also vessel permeability (33). The present study went further by estimating tumor blood flow and blood volume fraction from DCE-MRI during NACT, but still found no correlations. One possible explanation for the lack of correlation might be significant tissue heterogeneity in tumors; the parameters were estimated from the whole-tumor regions. Where possible, two smaller regions of interest (5×5 pixels) in each whole-tumor region were generated to reduce the likelihood of heterogeneity. It was assumed that these smaller regions would be more homogenous. However, no clear correlations were found in these smaller regions, and the data were observed to be more variable than the whole-tumor region, as reflected by the number of outliers and a wider range in the box plot scale (Figure 4, Supplementary Figure A1). An alternative method for future studies that might aid in selecting homogeneous tumor regions could be histogram analysis of pixel-wise IVIM and DCE-MRI parameter maps; however, the possibility of finding a homogeneous tumor region in the IVIM and DCE perfusion-related parameter maps to examine the correlation would require further investigation and validation.

Imprecision in the estimates of microvascular blood flow and blood volume fraction, in particular, is a potential issue that may have masked correlations between the IVIM and DCE-MRI perfusion parameters. A previous report recognized that the precision with which pseudo-diffusion is estimated is poor (34), and the estimate of blood volume fraction in another study was reported to be very imprecise (35), which was reflected in our calculated upper estimate of its repeatability (Appendix C, Supplementary Material). The estimation of blood volume fraction, against which the perfusion fraction derived from IVIM is compared, becomes difficult when tumor capillaries are excessively leaky (24). In this study, out of 108 DCE-MRI datasets, a one-compartment model was preferred in 13 cases, and an estimate of blood volume fraction and capillary permeability–surface area product was not possible in those 13.

It is also possible that IVIM and DCE-MRI reflect different underlying physiology. IVIM does not estimate perfusion in a

10.3389/fonc.2024.1356173

TABLE 2 Correlation between averaged DWI and DCE-MRI parameters from three MRI visits (Whole-tumor region).

Parameter		Tumor T_1	F_b	PS	v_e	v_b	v_d
ADC	**r**	**0.603****	0.026	0.305	**0.461***	-0.173	0.302
	P-value	**<0.001**	0.873	0.056	**0.004**	0.286	0.058
	N	**40**	40	40	**37**	40	40
D_t	**r**	**0.631****	0.014	0.266	**0.405***	-0.135	0.302
	P-value	**<0.001**	0.932	0.097	**0.013**	0.406	0.058
	N	**40**	40	40	**37**	40	40
D_p	**r**	-0.251	0.172	-0.051	-0.360	-0.006	-0.213
	P-value	0.118	0.289	0.755	0.029	0.971	0.187
	N	40	40	40	37	40	40
f	**r**	0.187	0.121	0.186	0.093	-0.144	0.079
	P-value	0.248	0.457	0.251	0.584	0.375	0.628
	N	40	40	40	37	40	40
$f \times D_p$	**r**	-0.020	0.143	0.071	-0.126	-0.041	0.001
	P-value	0.903	0.379	0.663	0.457	0.802	0.995
	N	40	40	40	37	40	40

r, correlation coefficient; N, sample siz;. ADC, apparent diffusion coefficient; D_t, tissue diffusion; D_p, pseudo-diffusion coefficient; f, perfused fraction; $f \times D_p$, microvascular blood flow; F_b, blood flow; PS, capillary permeability–surface area product; v_e, interstitial volume fraction; v_b, blood volume fraction; v_d, extracellular volume fraction.
Values in bold indicate significant correlation results: * $r \geq 0.4$ and $P<0.05$.
** $r \geq 0.4$ and $P<0.001$.

classical way but estimates flow in the direction of the diffusion encoding gradient, whereas DCE-MRI measures the delivery of blood and subsequent distribution of contrast agent in the tissue, on a different time scale (36). Furthermore, it has been suggested that a single pseudo-diffusion coefficient is insufficient to describe the complex diffusion properties of the vascular signal (37). The inconsistent patterns of response to treatment seen in the median values of perfusion fraction versus blood volume fraction and microvascular blood flow versus blood flow may support this suggestion (Figure 4).

In contrast, this study found moderate positive between-subject and within-subject repeated measures correlations between the diffusion parameters (ADC and D_t) and interstitial volume fraction, as well as a moderate positive within-subject repeated measures correlation between the diffusion parameters and extracellular volume fraction. These positive results are important, as this is the first time they have been observed in breast cancer (8), and support the current understanding of these imaging parameters. A positive between-subject correlation between ADC and interstitial volume fraction was previously determined in head and neck cancers (9) suggesting that these parameters are related to tissue microstructure. The ADC and D_t values reflect the diffusion of water molecules in tissue, which is affected by cellular density, membrane permeability and extracellular volume (7), and v_d is a direct measure of the extracellular volume fraction (24) while v_e is a parameter that reflects the volume fraction of the interstitial space within the tissue, which can be influenced by such factors as cellular density and extracellular matrix deposition. A prior study revealed that tumor cellularity is inversely proportional to v_e, v_d, and ADC values (38). Therefore, the observed between-subject correlation of the diffusion coefficients and interstitial space may suggest that breast tumors with a high cellular density tend to have a small interstitium and increased diffusion restriction, whereas tumors with a low cellular density tend to have a large interstitium and less diffusion restriction. The observed positive within-subject repeated measures correlations could result from the fact that ADC/D_t, v_e, and v_d exhibited similar patterns of change in response to treatment, wherein the values were increasing during the three MRI time-points (Figure 4).

Furthermore, a moderate positive between-subject correlation between the diffusion coefficients and tumor T_1 was observed in this study. Tumor T_1 measures tissue relaxation time, which can be affected by tissue water and fat content, macromolecule concentration and hydration state (39). Thus, this positive correlation may be because breast tumors with high cellular density and a small extracellular space have a decreased free-water content, resulting in low diffusion coefficient values and short tumor T_1 (12, 39).

The present study has some limitations. First, this study was performed on a limited sample size using a 1.5 T MRI scanner, which may limit the statistical power of the results. However, this is the first study that assesses both between-subject and within-subject repeated correlations between the perfusion parameters estimated by IVIM and DCE-MRI in a cohort of breast cancer patients undergoing

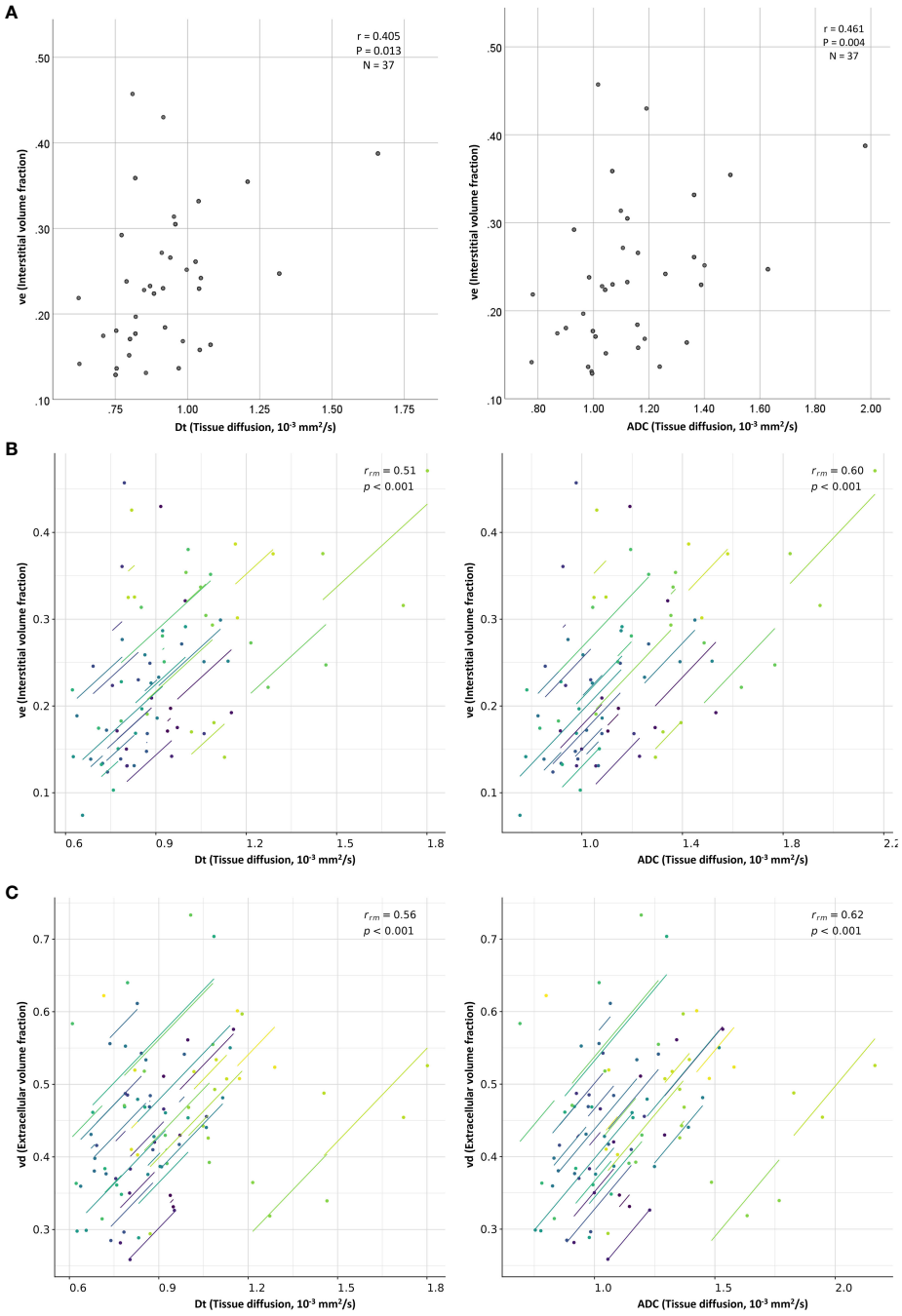


FIGURE 3
Scatter plots show moderate positive **(A)** between-subject and **(B)** within-subject repeated measures correlations between the diffusion coefficients (ADC and D_t) and the interstitial volume fraction (v_e), as well as moderate positive **(C)** within-subject repeated measures correlation between the diffusion coefficients (ADC and D_t) and the extracellular volume fraction (v_d). Each line in the scatter plots (**B**, **C**; repeated measures correlations) shows the fit for a single patient.

NACT with a primary focus on hypothesis generation rather than testing; therefore, the results can be used to direct future studies. Second, the DWI data were acquired with only 6 b-values, four of which were low (≤ 200 s/mm^2). In a clinical protocol, it is not practical to acquire DWI data with a large number of b-values. Nevertheless, the simulation study showed that using 6 b-values will not result in appreciably worse outcomes for most parameters, though the precision of microvascular blood flow was lower than with 12 b-values (details provided in Appendix B, Supplementary Material). Further, a previous study showed that a small number of b-values is not the main source of errors in IVIM parameter estimates. Intra-patient variability is significant; they found that the precision in the estimates of the IVIM parameters with only 4 b-values was better than the test-retest repeatability of those same parameters estimated

10.3389/fonc.2024.1356173

TABLE 3 Repeated measures correlations between DWI and DCE-MRI parameters estimated from Whole-tumor region.

Parameter		Tumor T_1	F_b	PS	v_e	v_b	v_d
ADC	r_{rm}	0.035	-0.361	-0.138	**0.597****	0.226	**0.619****
	df	67	67	54	**37**	54	**47**
	P-value	0.775	0.002	0.309	**<0.001**	0.094	**<0.001**
	95% CI	-0.18, 0.253	-0.605, 0.01	-0.452, 0.253	**0.203, 0.785**	-0.012, 0.432	**0.383, 0.82**
D_t	r_{rm}	0.043	-0.32	-0.045	**0.514****	0.165	**0.564****
	df	67	67	54	**37**	54	**47**
	P-value	0.724	0.007	0.741	**<0.001**	0.224	**<0.001**
	95% CI	-0.217, 0.279	-0.544, 0.036	-0.339, 0.312	**0.103, 0.716**	-0.052, 0.373	**0.305, 0.785**
D_p	r_{rm}	0.125	0.336	0.157	-0.127	-0.208	0.078
	df	65	65	53	37	53	46
	P-value	0.313	0.005	0.253	0.442	0.127	0.597
	95% CI	-0.08, 0.268	0.092, 0.502	-0.08, 0.397	-0.311, 0.074	-0.402, -0.02	-0.221, 0.304
f	r_{rm}	0.04	-0.182	-0.237	0.354	0.165	0.297
	df	65	65	53	37	53	46
	P-value	0.748	0.139	0.081	0.027	0.229	0.04
	95% CI	-0.201, 0.229	-0.423, 0.137	-0.477, 0.034	0.017, 0.583	-0.19, 0.418	0.068, 0.509
$f{\times}D_p$	r_{rm}	0.055	0.029	-0.059	0.215	-0.035	0.252
	df	65	65	53	37	53	46
	P-value	0.661	0.815	0.668	0.188	0.799	0.084
	95% CI	-0.178, 0.228	-0.169, 0.265	-0.222, 0.123	-0.066, 0.444	-0.261, 0.187	-0.023, 0.445

r_{rm}, repeated measures correlation coefficient; df, degrees of freedom; CI, confidence interval; ADC, apparent diffusion coefficient; D_t, tissue diffusion; D_p, pseudo-diffusion coefficient; f, perfused fraction; $f{\times}D_p$, microvascular blood flow; F_b, blood flow; PS, capillary permeability–surface area product; v_e, interstitial volume fraction; v_b, blood volume fraction; v_d, extracellular volume fraction.

Values in bold indicate significant correlation results: * $r_{rm} \geq 0.4$, P<0.05, and bootstrapped 95% CIs excluded zero.

** $r_{rm} \geq 0.4$, P<0.001 and bootstrapped 95% CIs excluded zero.

with 16 b-values (34). Third, a pixel-wise comparison of IVIM and DCE-MRI parameter maps was not performed in this study, although it might be valuable. Instead, the images were analyzed by following the recommended approaches of the International Breast Diffusion-Weighted Imaging Working Group (22), which included volumetric sampling and focused regions of interest (i.e., smaller single-slice regions on the darkest part of the ADC map). No correlations were observed between perfusion fraction versus blood volume fraction and microvascular blood flow versus blood flow in these smaller regions, but they showed more variability in the estimates instead (Supplementary Figure A1), suggesting that a pixel-wise analysis might yield similar outcomes. Fourth, rigid registration was employed for aligning the DCE and DWI images and this approach may not have been sufficient to correct DWI distortions. As such, the accuracy of spatial co-registration could have been affected, potentially influencing the findings reported, particularly in the smaller regions. Therefore, future work incorporating pixel-wise analysis following rigorous DWI and DCE-MRI image registration is needed to further investigate these relationships. Finally, the repeatability of the DWI and DCE-MRI parameters was not formally investigated. It was challenging to justify performing a repeated baseline DCE-MRI scan that required an additional injection of gadolinium contrast because the patients were due to undergo multiple NACT cycles and MRI scans. Instead, an upper estimate of the repeatability of the DWI and DCE-MRI parameters was calculated from a selection of baseline and cycle 1 studies (Appendix C, Supplementary Material).

In conclusion, this preliminary study investigated both between-subject and within-subject repeated measures correlations between DWI and DCE-MRI parameters in a cohort of patients with breast cancer imaged before and after one and three cycles of NACT. No statistically significant correlations were observed between the perfusion parameters estimated by IVIM (perfusion fraction and microvascular blood flow) and those estimated by DCE-MRI (blood flow and blood volume fraction). The two techniques may reflect different underlying physiology, and/or estimates of the IVIM and DCE-MRI parameters in the current study are largely imprecise. Therefore, care should be taken when interpreting the IVIM perfusion parameters as surrogates for those measured using DCE-MRI until their underlying pathophysiologic interpretation and relationship to the DCE-MRI perfusion parameters are elucidated by further research. However, the moderate positive within-subject repeated measures correlations found between the diffusion parameters and DCE-MRI measures of the tissue's interstitial and

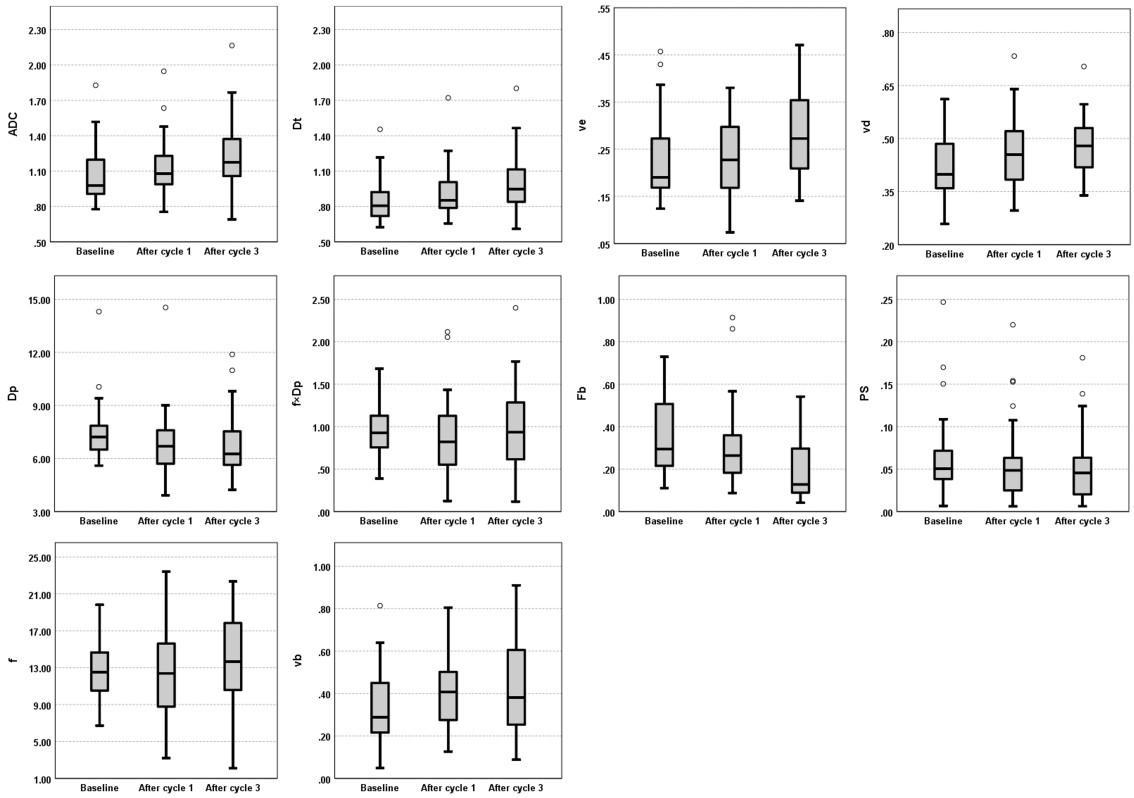


FIGURE 4
Evolution of DWI (ADC, D_t, D_p, f, and f×D_p) and DCE-MRI (F_b, v_b, PS, v_e, and v_d) parameters across the three MRI visits (Baseline, and after one and three cycles of NACT). Box plots illustrate the median and interquartile range values of all patients for whole-tumor region at each MRI visit. ADC, apparent diffusion coefficient; D_t, tissue diffusion; v_e, interstitial volume fraction; v_d, extracellular volume fraction; D_p, pseudo-diffusion coefficient; f×D_p, microvascular blood flow; F_b, blood flow; PS, capillary permeability–surface area product; f, perfused fraction; v_b, blood volume fraction.

extracellular volume fractions confirms the expectation that as these volumes increase, water diffusion increases.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Yorkshire & The Humber - Bradford Leeds Research Ethics Committee (Ethics approval: 15/YH/0246). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

ZA: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal Analysis, Data curation. SB: Writing – review & editing, Software, Methodology. DW: Writing – review & editing, Supervision, Funding acquisition. NS: Writing – review & editing, Funding acquisition. TD: Writing – review & editing, Supervision. DB: Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Data curation, Conceptualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Breast Cancer Now [award 2014MayPR241].

Acknowledgments

The study could not have been undertaken without the hard work and support of the following team members: Leonidas Georgiou, Andrew M. Hanby, Timothy J. Perren, David Dodwell, and Barbara J.G. Dall. The authors also gratefully acknowledge Jonathan Z. Bakdash and Laura R. Marusich for their generous contribution of time and expertise in providing comprehensive responses to inquiries regarding the rmcorr test.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1356173/full#supplementary-material

References

1. Organization WH. Breast Cancer (2023). Available online at: https://www.who.int/news-room/fact-sheets/detail/breast-cancer.

2. Kaufmann M, Von Minckwitz G, Mamounas EP, Cameron D, Carey LA, Cristofanilli M, et al. Recommendations from an international consensus conference on the current status and future of neoadjuvant systemic therapy in primary breast cancer. *Ann Surg Oncol.* (2012) 19:1508–16. doi: 10.1245/s10434-011-2108-2

3. Fowler AM, Mankoff DA, Joe BN. Imaging neoadjuvant therapy response in breast cancer. *Radiology*. (2017) 285:358–75. doi: 10.1148/radiol.2017170180

4. Shenoy H, Peter M, Masannat Y, Dall B, Dodwell D, Horgan KJSO. Practical advice on clinical decision making during neoadjuvant chemotherapy for primary breast cancer. *Surg Oncol*. (2009) 1:65–71. doi: 10.1016/j.suronc.2008.07.005

5. Georgiou L, Sharma N, Broadbent DA, Wilson DJ, Dall BJ, Gangi A, et al. Estimating breast tumor blood flow during neoadjuvant chemotherapy using interleaved high temporal and high spatial resolution mri. *Magn Reson Med*. (2018) 79:317–26. doi: 10.1002/mrm.26684

6. Mazhar SM, Shiehmorteza M, Kohl CA, Middleton MS, Sirlin CB. Nephrogenic systemic fibrosis in liver disease: A systematic review. *J Magn Reson Imaging*. (2009) 30:1313. doi: 10.1002/jmri.21983

7. Partridge S, Nissan N, Rahbar H, Kitsch A, Sigmund E. Diffusion-weighted breast mri: clinical applications and emerging techniques. *J Magn Reson Imaging*. (2017) 45:337–55. doi: 10.1002/jmri.v45.2

8. Arlinghaus LR, Li X, Rahman AR, Welch EB, Xu L, Gore JC, et al. On the relationship between the apparent diffusion coefficient and extravascular extracellular volume fraction in human breast cancer. *Magn Reson Imaging*. (2011) 29:630–8. doi: 10.1016/j.mri.2011.02.004

9. Han M, Kim SY, Lee SJ, Choi JW. The correlations between mri perfusion, diffusion parameters, and 18f-fdg pet metabolic parameters in primary head-and-neck cancer: A cross-sectional analysis in single institute. *Medicine*. (2015) 94 e2141. doi: 10.1097/MD.0000000000002141

10. Le Bihan D, Breton E, Lallemand D, Grenier P, Cabanis E, Laval-Jeantet M. Mr imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*. (1986) 161:401–7. doi: 10.1148/radiology.161.2.3763909

11. Le Bihan D, Turner R. The capillary network: A link between ivim and classical perfusion. *Magn Reson Med*. (1992) 27:171–8. doi: 10.1002/mrm.1910270116

12. Iima M, Kataoka M, Kanao S, Onishi N, Kawai M, Ohashi A, et al. Intravoxel incoherent motion and quantitative non-gaussian diffusion mr imaging: evaluation of the diagnostic and prognostic value of several markers of Malignant and benign breast lesions. *Radiology*. (2018) 287:432–41. doi: 10.1148/radiol.2017162853

13. Liu C, Liang C, Liu Z, Zhang S, Huang B. Intravoxel incoherent motion (Ivim) in evaluation of breast lesions: comparison with conventional dwi. *Eur J Radiol*. (2013) 82: e782–e9. doi: 10.1016/j.ejrad.2013.08.006

14. Almutlaq ZM, Wilson DJ, Bacon SE, Sharma N, Stephens S, Dondo T, et al. Evaluation of monoexponential, stretched-exponential and intravoxel incoherent motion mri diffusion models in early response monitoring to neoadjuvant chemotherapy in patients with breast cancer—a preliminary study. *J Magn Reson Imaging*. (2022) 56:1079–88. doi: 10.1002/jmri.28113

15. Bedair R, Priest A, Patterson A, McLean M, Graves M, Manavaki R, et al. Assessment of early treatment response to neoadjuvant chemotherapy in breast cancer using non-mono-exponential diffusion models: A feasibility study comparing the baseline and mid-treatment mri examinations. *Eur Radiol*. (2017) 27:2726–36. doi: 10.1007/s00330-016-4630-x

16. Che S, Zhao X, Yanghan O, Li J, Wang M, Wu B, et al. Role of the intravoxel incoherent motion diffusion weighted imaging in the pre-treatment prediction and early response monitoring to neoadjuvant chemotherapy in locally advanced breast cancer. *Medicine*. (2016) 95:1–12. doi: 10.1097/MD.0000000000002420

17. Liu C, Wang K, Chan Q, Liu Z, Zhang J, He H, et al. Intravoxel incoherent motion mr imaging for breast lesions: comparison and correlation with pharmacokinetic evaluation from dynamic contrast-enhanced mr imaging. *Eur Radiol*. (2016) 26:3888–98. doi: 10.1007/s00330-016-4241-6

18. Jiang L, Lu X, Hua B, Gao J, Zheng D, Zhou Y. Intravoxel incoherent motion diffusion-weighted imaging versus dynamic contrast-enhanced magnetic resonance imaging: comparison of the diagnostic performance of perfusion-related parameters in breast. *J Comput Assist Tomogr*. (2018) 42:6–11. doi: 10.1097/RCT.0000000000000661

19. Li K, Machireddy A, Tudorica A, Moloney B, Oh KY, Jafarian N, et al. Discrimination of Malignant and benign breast lesions using quantitative multiparametric mri: A preliminary study. *Tomography*. (2020) 6:148–59. doi: 10.18383/j.tom.2019.00028

20. Partridge SC, Zhang Z, Newitt DC, Gibbs JE, Chenevert TL, Rosen MA, et al. Diffusion-weighted mri findings predict pathologic response in neoadjuvant treatment of breast cancer: the acrin 6698 multicenter trial. *Radiology*. (2018) 289:618–27. doi: 10.1148/radiol.2018180273

21. Georgiou L, Wilson DJ, Sharma N, Perren TJ, Buckley DL. A functional form for a representative individual arterial input function measured from a population using high temporal resolution dce mri. *Magn Reson Med*. (2019) 81:1955–63. doi: 10.1002/mrm.27524

22. Baltzer P, Mann RM, Iima M, Sigmund EE, Clauser P, Gilbert FJ, et al. Diffusion-weighted imaging of the breast-a consensus and mission statement from the eusobi international breast diffusion-weighted imaging working group. *Eur Radiol*. (2020) 30:1436–50. doi: 10.1007/s00330-019-06510-3

23. Cron GO, Santyr G, Kelcz F. Accurate and rapid quantitative dynamic contrast-enhanced breast mr imaging using spoiled gradient-recalled echoes and bookend T1 measurements. *Magn Reson Med*. (1999) 42:746–53. doi: 10.1002/(ISSN)1522-2594

24. Sourbron SP, Buckley DL. Classic models for dynamic contrast-enhanced mri. *NMR BioMed*. (2013) 26:1004–27. doi: 10.1002/nbm.2940

25. Buckley DL. Shutter-speed dynamic contrast-enhanced mri: is it fit for purpose? *Magn Reson Med*. (2019) 81:976–88. doi: 10.1002/mrm.27456

26. Glatting G, Kletting P, Reske SN, Hohl K, Ring C. Choosing the optimal fit function: comparison of the akaike information criterion and the F-test. *Med Phys*. (2007) 34:4285–92. doi: 10.1118/1.2794176

27. Iima M, Partridge SC, Le Bihan D. Six dwi questions you always wanted to know but were afraid to ask: clinical relevance for breast diffusion mri. *Eur Radiol*. (2020) 30:2561–70. doi: 10.1007/s00330-019-06648-0

28. Suo S, Lin N, Wang H, Zhang L, Wang R, Zhang S, et al. Intravoxel incoherent motion diffusion-weighted mr imaging of breast cancer at 3.0 tesla: comparison of different curve-fitting methods. *J Magn Reson Imaging*. (2015) 42:362–70. doi: 10.1002/jmri.v42.2

29. Bland JM, Altman DG. Calculating correlation coefficients with repeated observations: part 2—Correlation between subjects. *BMJ*. (1995) 310:633. doi: 10.1136/bmj.310.6980.633

30. Overholser BR, Sowinski KM. Biostatistics Primer: Part 2. *Nutr Clin Pract*. (2008) 23(1):76–84. doi: 10.1177/011542650802300176

31. Bakdash JZ, Marusich LR. Repeated measures correlation. *Front Psychol*. (2017) 8:456. doi: 10.3389/fpsyg.2017.00456

32. Marusich LR, Bakdash JZ. Rmcorrshiny: A web and standalone application for repeated measures correlation [Version 1; peer review: 2 approved]. *F1000Research*. (2021) 10:697. doi: 10.12688/f1000research

33. Sourbron SP, Buckley DL. On the scope and interpretation of the tofts models for dce-mri. *Magn Reson Med*. (2011) 66:735–45. doi: 10.1002/mrm.22861

34. Dyvorne H, Jajamovich G, Kakite S, Kuehn B, Taouli B. Intravoxel incoherent motion diffusion imaging of the liver: optimal B-value subsampling and impact on

parameter precision and reproducibility. *Eur J Radiol*. (2014) 83:2109–13. doi: 10.1016/j.ejrad.2014.09.003

35. Kershaw LE, Hutchinson CE, Buckley DL. Benign prostatic hyperplasia: evaluation of T1, T2, and microvascular characteristics with T1-weighted dynamic contrast-enhanced mri. *J Magn Reson Imaging*. (2009) 29:641–8. doi: 10.1002/jmri.21674

36. Henkelman RM. Does ivim measure classical perfusion? *Magn Reson Med*. (1990) 16:470–5. doi: 10.1002/mrm.1910160313

37. Zhang X, Ingo C, Teeuwisse WM, Chen Z, van Osch MJ. Comparison of perfusion signal acquired by arterial spin labeling–prepared intravoxel incoherent motion (Ivim) mri and conventional ivim mri to unravel the origin of the ivim signal. *Magn Reson Med*. (2018) 79:723–9. doi: 10.1002/mrm.26723

38. Aryal MP, Nagaraja TN, Keenan KA, Bagher-Ebadian H, Panda S, Brown SL, et al. Dynamic contrast enhanced mri parameters and tumor cellularity in a rat model of cerebral glioma at 7 T. *Magn Reson Med*. (2014) 71:2206–14. doi: 10.1002/mrm.24873

39. Meng L, Zhao X, Guo J, Lu L, Cheng M, Xing Q, et al. Evaluation of the differentiation of benign and Malignant breast lesions using synthetic relaxometry and the kaiser score. *Front Oncol*. (2022) 12. doi: 10.3389/fonc.2022.964078

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 06 September 2024
DOI 10.3389/fonc.2024.1399296

OPEN ACCESS

EDITED BY
Luca Nicosia,
University of Insubria, Italy

REVIEWED BY
Mehmet Dalmis,
Meta, Netherlands
Arutselvan Natarajan,
Stanford University, United States
Ujjwal Agarwal,
Tata Memorial Hospital, India

*CORRESPONDENCE
Shijie Wang
wangshj37@mail3.sysu.edu.cn

†These authors have contributed equally to this work

RECEIVED 11 March 2024
ACCEPTED 19 August 2024
PUBLISHED 06 September 2024

CITATION
Li Y, Li C, Yang T, Chen L, Huang M, Yang L, Zhou S, Liu H, Xia J and Wang S (2024) Multiview deep learning networks based on automated breast volume scanner images for identifying breast cancer in BI-RADS 4. *Front. Oncol.* 14:1399296. doi: 10.3389/fonc.2024.1399296

Multiview deep learning networks based on automated breast volume scanner images for identifying breast cancer in BI-RADS 4

Yini Li[1†], Cao Li[2†], Tao Yang[1], Lingzhi Chen[1], Mingquan Huang[3], Lu Yang[2], Shuxian Zhou[4], Huaqing Liu[4], Jizhu Xia[1] and Shijie Wang[1*]

[1]Department of Ultrasound, The Affiliated Hospital of Southwest Medical University, Sichuan, China, [2]Department of Radiology, The Affiliated Hospital of Southwest Medical University, Sichuan, China, [3]Department of Breast Surgery, The Affiliated Hospital of Southwest Medical University, Sichuan, China, [4]Artificial Intelligence Innovation Center, Research Institute of Tsinghua, Guangdong, China

Objectives: To develop and validate a deep learning (DL) based automatic segmentation and classification system to classify benign and malignant BI-RADS 4 lesions imaged with ABVS.

Methods: From May to December 2020, patients with BI-RADS 4 lesions from Centre 1 and Centre 2 were retrospectively enrolled and divided into a training set (Centre 1) and an independent test set (Centre 2). All included patients underwent an ABVS examination within one week before the biopsy. A two-stage DL framework consisting of an automatic segmentation module and an automatic classification module was developed. The preprocessed ABVS images were input into the segmentation module for BI-RADS 4 lesion segmentation. The classification model was constructed to extract features and output the probability of malignancy. The diagnostic performances among different ABVS views (axial, sagittal, coronal, and multi-view) and DL architectures (Inception-v3, ResNet 50, and MobileNet) were compared.

Results: A total of 251 BI-RADS 4 lesions from 216 patients were included (178 in the training set and 73 in the independent test set). The average Dice coefficient, precision, and recall of the segmentation module in the test set were 0.817 ± 0.142, 0.903 ± 0.183, and 0.886 ± 0.187, respectively. The DL model based on multiview ABVS images and Inception-v3 achieved the best performance, with an AUC, sensitivity, specificity, PPV, and NPV of 0.949 (95% CI: 0.945-0.953), 82.14%, 95.56%, 92.00%, and 89.58%, respectively, in the test set.

Abbreviations: BI-RADS, Breast Imaging Reporting and Data System; ABVS, automated breast volume scanner; DL, deep learning; US, ultrasound; AI, artificial intelligence; MRI, magnetic resonance imaging; CT, computed tomography; AUC, area under the curve; CNN, convolutional neural network; ROC, receiver operating characteristic; PPV, positive predictive value; NPV, negative predictive value; DC, Dice coefficient; Grad-CAM, gradient-weighted class activation mapping.

Conclusions: The developed multiview DL model enables automatic segmentation and classification of BI-RADS 4 lesions in ABVS images.

KEYWORDS

BI-RADS 4, deep learning, breast cancer, automated breast ultrasound, segmentation

1 Introduction

Breast cancer has become the most prevalent cancer worldwide, with 2.3 million new cases resulting in 665,684 deaths in 2022 (1). Accurate identification and timely treatment are effective measures to reduce its mortality. Breast Imaging Reporting and Data System (BI-RADS) category 4 lesions (2) are suspected to be malignant lesions (2%~95% likelihood) and are recommended for biopsies, which results in more than 67.0% of benign lesions receiving biopsies (3–6). This may lead to unnecessary anxiety and invasive examination-related complications, such as pain, infection, and needle track seeding, in patients as well as increase the burden to the healthcare system (7). Therefore, a noninvasive method for identifying malignant BI-RADS 4 lesions and reducing unnecessary biopsies is an urgent issue in current precision medicine.

Ultrasound (US) is not inferior to mammography for screening for breast cancer and has a sensitivity of up to 90% in dense breasts with safety and low cost (8, 9). The automated breast volume scanner (ABVS) is a novel breast ultrasound imaging technique that overcomes many of the limitations of traditional US, provides a three-dimensional (3D) representation of breast tissue, and allows image reformatting in three planes (axial, sagittal and coronal) (10). The ABVS' unique coronal images provide an intuitive view of the lesions and their relationships with neighboring catheters and surrounding tissues. The retraction phenomenon, characterized by a perinodal stripe of hypoechoic and hyperechoic radial extension, is a unique sign on the coronal plane for malignant breast tumors with a high specificity (91.1%~100%) (11, 12). However, the large amount of ABVS image data is a significant challenge for radiologists.

Deep learning (DL) is a subfield of artificial intelligence (AI), and its emergence has increased interest in automated detection and diagnostic tools in medicine (13). DL has achieved state-of-the-art performance in feature recognition and classification in several modalities, including magnetic resonance imaging (MRI), computed tomography (CT), X-ray and US (14–17). Recent studies have shown that DL methods using ABVS images also have enormous potential in breast cancer (18–20). Wang et al. (18) proposed a DL method that adopted a modified Inception-v3 architecture to extract effective features from ABVS images to distinguish between benign and malignant breast lesions with an area under the curve (AUC), sensitivity, and specificity of 0.945, 0.886, and 0.876, respectively. However, most of these studies were designed as proof-of-concept or technical feasibility studies without a thorough external validation of real-world clinical performance (19, 20). To our knowledge, no studies have investigated the use of DL methods based on ABVS images to distinguish between benign and malignant BI-RADS 4 lesions.

Therefore, we attempted to develop a DL model based on ABVS images with automatic segmentation and classification capabilities, and to explore its performance in identifying benign and malignant BI-RADS 4 lesions and in reducing unnecessary biopsies. In addition, since ABVS images can be visualized in axial, sagittal and coronal views, we further compared the DL models based on the use of single or multiple views.

2 Materials and methods

This retrospective study was approved by the Institutional Review Board of the Affiliated Hospital of Southwest Medical University (KY2020163) and was conducted following the Declaration of Helsinki guidelines. All participating subjects were informed and voluntarily signed informed consent forms.

2.1 Patients and data collection

From 1 May to 31 December, 2020, consecutive patients with BI-RADS 4 lesions on US who were scheduled for biopsies at the Affiliated Hospital of Southwest Medical University (Centre 1, the training set) and Guangdong Provincial Hospital of Traditional Chinese Medicine (Centre 2, the independent test set) were invited to participate in this study. Further selection was performed according to the following inclusion and exclusion criteria.

The inclusion criteria were as follows: (1) age≥18 years; (2) BI-RADS 4 lesions identified following the 2013 edition of the BI-RADS guidelines (21) by two senior radiologists (>10 years of breast US experience) at both centers; and (3) completion of the ABVS examination within one week before biopsy. The exclusion criteria were as follows: (1) patients who were breastfeeding or had mastitis or breaks in the affected breast; (2) patients who had undergone previous invasive procedures for the lesion; (3) patients with poor-quality images; and (4) patients who lacked definitive pathologic findings. Patients with more than one BI-RADS 4 lesion were included separately.

Clinical data included age, menopausal status, history of oral contraceptive use and smoking history, alcohol consumption level, and family history of breast or ovarian cancer. The patient's breast density was classified as type A-D according to the mammographic BI-RADS guidelines. The characteristics of the lesions, including the lesion size, location (left or right), shape (regular or irregular), orientation (parallel or nonparallel), posterior echogenicity (enhancement, shadowing, mixed pattern, or absence of posterior echogenicity), internal echogenicity (hypoechoic, hyperechoic, or mixed echogenicity), and calcification (present or absent), were recorded.

2.2 ABVS examinations

All ABVS examinations were performed by the Acuson S2000 ABVS (Siemens, Germany) ultrasound systems with the 14L5BV probe (5–14 MHz) by two technicians (with 6 months of ABVS training experience). For more details on the ABVS examination, see Kim et al (22). After the examination, axial ABVS images were sent to a dedicated workstation, and the sagittal and coronal images were reconstructed automatically. Finally, the axial, sagittal, and coronal ABVS images showing the largest lesions were selected for further segmentation and classification. An example is shown in Figure 1.

Within one week after the ABVS examination, a US-guided core-needle biopsy was performed by experienced US doctors. In accordance with the standard biopsy procedure, four to eight samples per lesion were acquired via an automatic biopsy gun with a 14G or 16G needle. The specimens were analyzed and diagnosed by breast pathologists (>10 years of experience), according to the World Health Organization's standards for breast tumor classification (23). For lesions with unclear diagnoses by puncture, histopathologic diagnosis after surgical removal was used as the reference standard.

2.3 DL Framework and models

Centers 1 and 2 were divided into a training set and an independent test set, respectively. We utilized five-fold cross-validation on the training set to optimize the parameters of the models and guide the choice of hyperparameters. The test set was used to evaluate the final model performance independently. A two-stage DL framework consisting of an automatic segmentation module and an automatic classification module was developed. First, the preprocessed ABVS images were input into the automatic segmentation module for BI-RADS 4 lesion segmentation. Patches were created as the input to the classifier. The classification model was subsequently constructed via convolutional neural networks (CNNs) to automatically extract the features of the lesions and output the probability of malignancy. The overall process is described in detail below and the whole pipeline of the DL model is shown in Figure 2. Finally, we visualized and analyzed the prediction results of the DL model.

2.3.1 Image preprocessing and automatic segmentation module

Histogram equalization and median filtering were used to remove noise and enhance the images. The black boxes in the ABVS images were cropped using the Sobel operator (24). Online data augmentation was performed for the ABVS images in the training set during the training period. The augmented image pixels were normalized and input into the ImageNet dataset for pretraining.

The DeepLab-V3 algorithm introduced by Google was used to build the automatic image segmentation module. DeepLab-V3 uses the atrous spatial pyramid pooling (ASPP) structure to expand the receptive field, mining context information, and the improved Xception module to reduce the number of parameters and achieve the best effect of the current segmentation network.

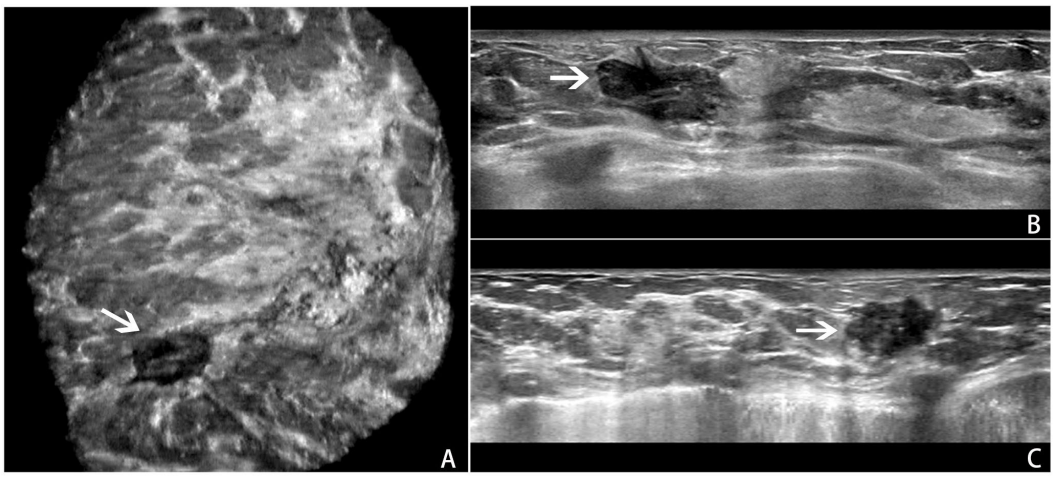


FIGURE 1
An example of a BI-RADS 4 lesion on ABVS images.ABVS images of the largest sections of a lesion in the coronal **(A)**, axial **(B)**, and sagittal **(C)** planes.

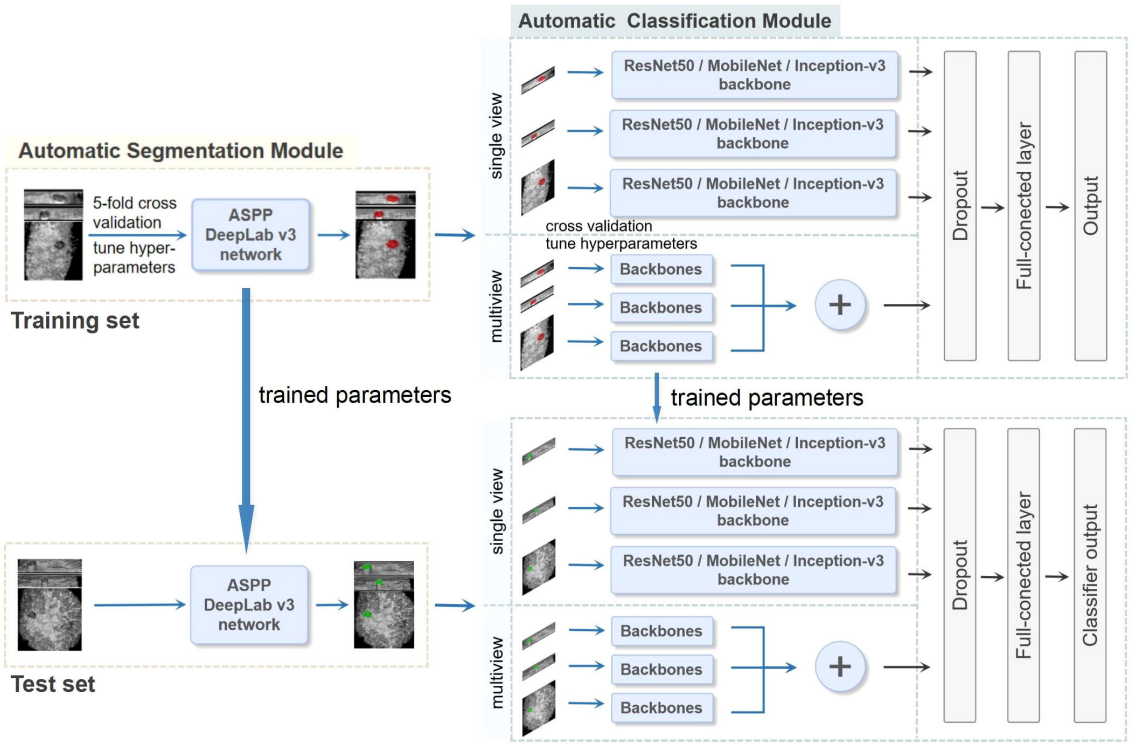


FIGURE 2
The whole pipeline of the deep learning (DL) model. The illustration shows the image input and the main processing stages for the two-stage DL framework model, which consists of an automatic segmentation module and an automatic classification module. The preprocessed ABVS images were used as input for the segmentation module for segmenting lesions. The classification module was constructed on the basis of single-view (axial, sagittal, and coronal) and multiview (combined axial, sagittal, and coronal) images, as well as different backbone networks (ResNet50, MobileNet and Inception-v3), and outputs the probability of malignancy. In the multiview model, each view of the input images corresponds to a backbone network branch, and three branches are concatenated to form the total feature vector. Five-fold cross-validation was utilized on the training set to choose the hyperparameters. The test set was used to evaluate the final performance.

2.3.2 Automatic classification module

The segmented images of the lesion and its surrounding area were as patches to input to the classification module to extract features and automatically output the probability of malignancy. For the reasons that manually labelling masks has a certain degree of subjectivity; the segmentation results of the segmentation model also have certain biases; and the differences between the lesion area and nearby normal tissues may help AI classify more accurately. To construct the optimal DL model, we explored the performances of CNN models based on single-view (axial, sagittal, and coronal) and multiview (combined axial, sagittal, and coronal) images, as well as different backbone networks (ResNet50, MobileNet, and Inception-v3) in differentiating benign and malignant BI-RADS 4 lesions. Transfer learning was applied to ensure a strong feature extraction capability. Because of the limited number of samples, pretrained knowledge was effectively applied to a specific task from a mega database such as ImageNet, and the model was then retrained using a small amount of data, which could achieve satisfactory results (25). Each model was fine-tuned on the dataset of ABVS images to reduce overfitting. The convolutional structure was used as the backbone network, consisting of multiple convolutional layers, average pooling layers, and convolutional modules in series for feature extraction. In the multiview models, each view of the input images corresponds to a backbone network branch, and three branches are concatenated to form the total feature vector. A CNN framework example of Inception-v3 is shown in Figure 3. A dropout layer (with deactivation rate of 0.5) was added behind the vector to mitigate overfitting. Finally, the fully connected layer was normalized to output the probability of malignancy of BI-RADS 4 lesions (a cut-off value of 50%).

The image preprocessing methods and DL algorithms with the parameters and software settings are detailed in Supplementary File 1.

2.3.3 Testing and visualization of the DL model

The fine-tuned parameters were used in the segmentation and classification models of the independent test set to evaluate the effectiveness and final performance of these models. The results were analyzed and assessed by the area under the receiver operating characteristic (ROC) curve. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated at the maximum Youden index. The performance of the automatic segmentation network was evaluated via the Dice coefficient (DC). Additionally, we set a decision point in the ROC curve based on the final model where sensitivity is 100% to evaluate the value in reducing the unnecessary biopsies and this would allow no lesions to be missed.

Gradient-weighted class activation mapping (Grad-CAM) was used on the final convolutional layer of the classification model to

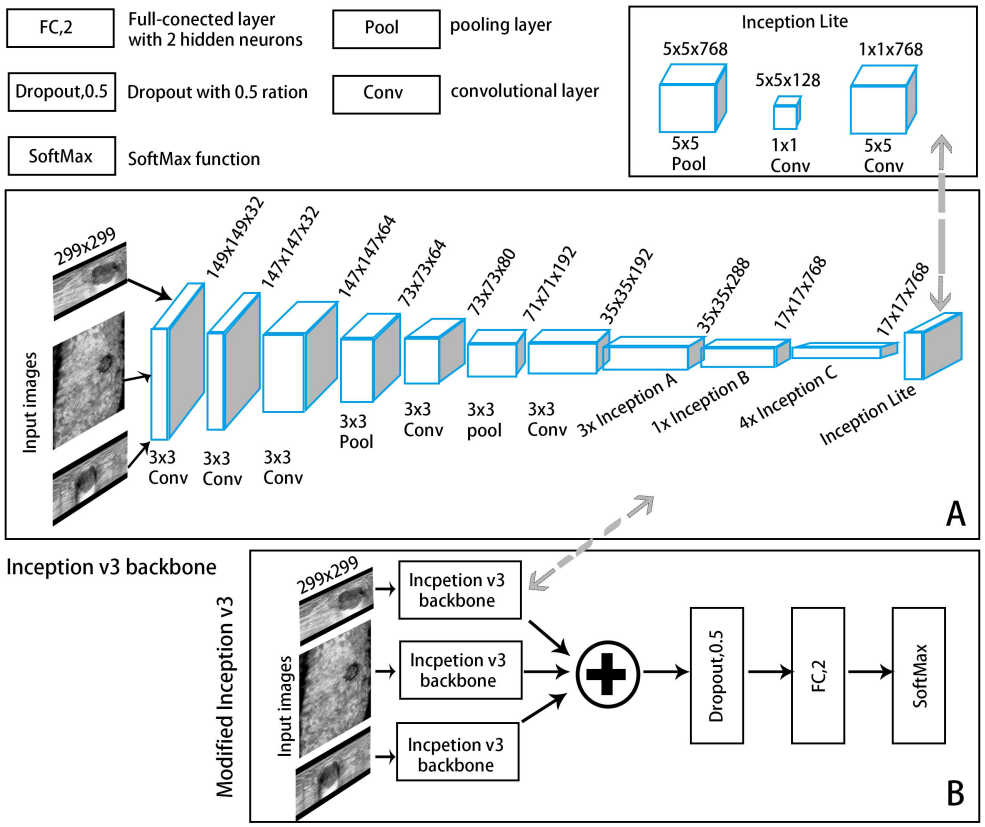


FIGURE 3
A convolutional neural network (CNN) framework example of an automatic classification model with Inception-v3 as the backbone network. **(A)** The backbone network: The input ABVS image in three views (axial, sagittal, and coronal) passed six convolutional layers and one average pooling layer, followed by three Inception A modules, one Inception B module, four Inception C modules and one Inception Lite module as defined in this study. The Inception Lite module consists of an average pooling layer in tandem with two convolutional modules of different kernel sizes. **(B)** The modified version of Inception-v3.

visualize the extent of each region on the ABVS image that contributed to identifying malignant BI-RADS 4 lesions. The critical areas predicted by the model are highlighted.

2.4 Statistical analysis

IBM SPSS Statistics (version 26.0, IBM Corp., USA) and Python software (version 3.6.8, https://www.python.org/) were used for the statistical analysis. SPSS software was used to analyze the differences between the training and test sets and between benign and malignant lesions. Continuous variables (age and tumor size) were compared via t-tests. Categorical variables (breast density, BI-RADS 4 subclasses, and family history of breast cancer) were compared via the chi-square test.

The DC, recall, and precision were introduced to evaluate the automatic segmentation performance objectively. ROC curves were constructed to assess the classification performance and to calculate the sensitivity, specificity, PPV, NPV, and AUC. The AUCs were compared via the DeLong test. All the statistical calculations were performed with 95% confidence intervals (95% CIs). All tests were two-sided, and $P<0.05$ was considered statistically significant.

3 Results

3.1 Patient characteristics

A total of 251 BI-RADS 4 lesions in 216 patients from two centers were included. The flow chart is shown in Figure 4. Among them, 178 lesions from 157 patients (mean age 49.0 ± 11.8 years) at Centre 1 were included in the training set, and 73 lesions from 59 patients (average age 46.8 ± 10.5 years) at Centre 2 were included in the independent test set. The proportions of malignant lesions between the two sets were not significantly different (45.5% *vs.* 39.7%, $P=0.402$), and there were no significant differences in patient age, lesion size, lesion location, BI-RADS 4 subclassifications, breast density, or family history (Table 1).

3.2 Performance of the automatic segmentation module

The Dice coefficient curves (Supplementary File 2) for assessing segmentation performance revealed 474 (88.7%) ABVS images with DCs greater than 0.90 in the training set and 165 (75.3%) ABVS

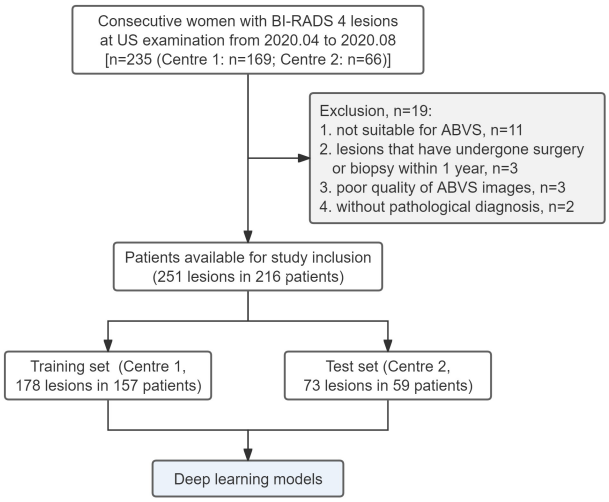


FIGURE 4
Flow chart of the study.

images in the test set. Example plots are provided in Supplementary File 3. The automatic segmentation model has the best segmentation performance in axial views with DCs, recall rates, and precisions of 0.908 ± 0.077, 0.996 ± 0.067, and 0.974 ± 0.123 in the training set and 0.890 ± 0.152, 0.972 ± 0.167, and 0.948 ± 0.198 in the test set, respectively. Among all the views, the segmentation module displayed the worst performance in the coronal views, with DC, recall and precision values of 0.784 ± 0.120, 0.825 ± 0.188, and 0.825 ± 0.188, respectively, which is still a satisfactory result. The detailed segmentation statistics of the different ABVS views in the two sets are shown in Table 2.

3.3 Performance of the automatic classification module

The automatic classification of BI-RADS 4 lesions was performed after automatically segmenting the lesions.

TABLE 1 Baseline data of the benign and malignant BI-RADS 4 lesions in the training and test sets.

Characteristics	Training set (n=178)			Test set (n=45)			*P* *
	Malignant (n=81)	Benign (n=97)	*P*	Malignant (n=29)	Benign (n=44)	*P*	
Age (years, ±)	54.6 ± 12.5	44.1 ± 9.9	<0.001	57.5 ± 9.4	45.4 ± 11.7	<0.001	0.229
Lesion size (cm, ±)*	2.3 ± 0.9	1.6 ± 0.9	<0.001	2.5 ± 1.2	1.6 ± 0.7	0.005	0.549
BI-RADS 4 category (n, %)							
4a	11(13.3%)	73(76.0%)	<0.001	2(6.9%)	35(79.5%)	<0.001	0.593
4b	16(19.3%)	20(20.0%)		4(13.8%)	9(20.5%)		
4c	54(67.4%)	4(4.0%)		23(79.3%)	0(0.0%)		
Breast density (n, %)							
A	11(13.2%)	4(4.0%)	<0.001	4(13.8%)	0(0.0%)	0.319	0.878
B	38(45.8%)	26(27.0%)		12(41.4%)	17(38.6%)		
C	26(45.8)	42(44.0%)		9(31.0%)	20(45.5%)		
D	6(7.2%)	25(25.0%)		4(13.8%)	7(15.9%)		
Menopausal status (n, %)							
Premenopausal	33(41.0%)	73(74.0%)	<0.001	6(20.7%)	29(65.9%)	0.010	0.054
Postmenopausal	48(59.0%)	24(26.0%)		23(79.3%)	15(34.1%)		
Family history (n, %)							
Yes	8(9.6%)	8(8.0%)	0.705	6(20.7%)	13(29.5%)	0.793	0.153
No	73(90.4%)	89(92.0%)		23(79.3%)	31(70.5%)		
Location of lesion (n, %)							
Left	51(62.7%)	52(55.05)	0.208	15(51.7%)	26(59.1%)	0.302	0.992
Right	30(37.3%)	45(45.0%)		14(48.2%)	18(40.9%)		

*P values between the training set and the test set.
Lesion size was defined as the maximum diameter on ABVS images. Family history referred to breast or ovarian cancer in first-degree relatives. The differences in characteristic variables (age and lesion size) between the two cohorts were compared via two-sample t-tests, whereas chi-square tests were conducted on the other variables. P<0.05.
BI-RADS, Breast Imaging Reporting and Data System.

Among the DL models on the various views, the multiview models had better classification performance than the single-view models in different backbone networks (ResNet50, MobileNet, and Inception-v3). Moreover, all three single-view models and the multiview model achieved the best classification performance in the Inception-v3 network on both sets. The statistics for the training set are shown in Supplementary File 4, and the performance results for the test set are shown in Table 3, Figure 5. Among them, the multiview model with the Inception-v3 backbone had the best performance, with an AUC, sensitivity, specificity, PPV, and NPV of 0.949 (95% CI: 0.945–0.953), 82.14%, 95.56%, 92.00%, and 89.58%, respectively. However, the coronal single-view model based on ResNet50 had the worst classification performance, with an AUC, sensitivity, specificity, PPV, and NPV of 0.807 (95% CI: 0.779–0.836), 85.71%, 57.78%, 55.81%, and 86.67%, respectively.

3.4 Value in reducing unnecessary biopsies and visualizations

The confusion matrix of all the DL models with the test set is shown in Figure 6. The Inception-v3-based multiview DL model performed the best, with a missed diagnosis rate and misdiagnosis rate of 17.85% (5/28) and 4.44% (2/45), and with the unnecessary biopsy rate reducing from 61.64% (45/73) to 8.00% (2/25) compared to the conventional US. The sagittal single-view and multiview models based on the MobileNet network achieved similar performance, with missed diagnosis rates and misdiagnosis rates of 21.43% (6/28) and 2.22% (1/45) on the sagittal view, and 17.85% (5/28) and 6.67% (3/45) on the multiview, respectively. With a decision point of 100% sensitivity in the ROC curve based on multiview Inception-v3 model, the specificity, PPV, and NPV were 58.1%, 58.9%, and 100%, respectively (The confusion matrix is shown in Supplementary File 5). And the unnecessary biopsy rate of it is 40.42% (19/47), which is 21.22% lower than conventional ultrasound (61.64%, 45/73) without missing any malignant lesions.

TABLE 2 Automatic segmentation results of different ABVS views in the training and test sets.

Clusters	Dice coefficient (mean ± SD)	Recall (mean ± SD)	Precision (mean ± SD)
The training set (n=534)	0.874 ± 0.173	0.911 ± 0.185	0.893 ± 0.195
Coronal view (n=178)	0.804 ± 0.121	0.894 ± 0.207	0.883 ± 0.216
Sagittal view (n=178)	0.824 ± 0.154	0.926 ± 0.163	0.905 ± 0.177
Axial view (n=178)	**0.908 ± 0.077**	**0.996 ± 0.067**	**0.974 ± 0.123**
The test set (n=219)	0.817 ± 0.142	0.903 ± 0.183	0.886 ± 0.187
Coronal view (n=73)	0.784 ± 0.120	0.825 ± 0.188	0.825 ± 0.188
Sagittal view (n=73)	0.801 ± 0.119	0.883 ± 0.157	0.888 ± 0.165
Axial view (n=73)	**0.890 ± 0.152**	**0.972 ± 0.167**	**0.948 ± 0.198**

SD refers to the standard deviation, and the bolded portion is the group with the best indicator.

The saliency map highlighted the lesion location and surrounding region, both in benign and malignant lesions (Figure 7). This finding indicated that the multiview DL model focused on the lesion itself and surrounding structures when categorizing BI-RADS 4 lesions.

4 Discussion

In this study, we developed an ABVS-based DL model with automatic segmentation and classification capabilities to explore its diagnostic performance in single-view and multiview images for identifying breast cancer in BI-RADS 4 lesions. We found that our DL model can accurately segment multiple views of ABVS images and further differentiate benign and malignant BI-RADS 4 lesions, which could reduce unnecessary invasive biopsies.

DL, a technique used in artificial intelligence, has achieved significant advances in automatic medical image analysis of breast cancer through CNNs. In addition to segmenting (26) and categorizing (27) various modalities of ultrasound images of breast cancer, DL can also predict metastasis (28) and patient prognosis (29). The BI-RADS 4 lesion is the watershed for whether to perform a biopsy, with a 5%–98% likelihood of being benign (3). Therefore, accurately differentiating the benign and malignant natures of BI-RADS 4 lesions is the key to minimizing noninvasive manipulation of breast lumps and is a pressing issue. Therefore, we used a deep learning approach to solve this problem noninvasively. To our knowledge, the development of ABVS-based DL models for the automatic segmentation and classification of BI-RADS 4 lesions, as well as the application of such an approach for reducing the possibility of biopsy, has not been reported.

This study used the DeepLab-V3 model to segment BI-RADS 4 lesions automatically, and the high DC values reflected its powerful segmentation performance. The segmentation effectiveness was the worst in the coronal plane. This may be because the artefacts caused by the nipple are extremely similar to the echogenicity of the lesion in the coronal plane, and the adipose tissue in the breast, which is morphologically similar to some breast nodules, is also in a restricted distribution in this plane. The model achieved the best segmentation performance in the axial single section, which is consistent with recent research results (30). Therefore, the DeepLab-V3-based segmentation module actually has excellent segmentation efficacy, self-learning ability, and self-adaptation for ABVS image segmentation (31, 32). Moreover, high-quality automatic segmentation lays a foundation for the subsequent standardization of feature extraction and classification accuracy (33).

ABVS can provide 3D images and reconstruct the images to axial, sagittal, and coronal views. Thus, we explored the

TABLE 3 Diagnostic performance of the single-view and multiview models based on different backbone networks (ResNet50, MobileNet, and Inception-v3) on the test set.

Backbones	View	AUC (95%CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
ResNet50	Axial view	0.880 (0.857-0.903)	71.43	91.11	83.33	83.67
	Sagittal view	0.898 (0.880-0.915)	82.14	82.22	74.19	88.10
	Coronal view	0.807 (0.779-0.836)	85.71	57.78	55.81	86.67
	Multiview *	**0.922 (0.905-0.939)**	**82.14**	**84.44**	**76.67**	**88.37**
MobileNet	Axial view	0.909 (0.886-0.931)	78.57	93.33	88.00	87.50
	Sagittal view	0.910 (0.890-0.930)	78.57	97.78	95.65	88.00
	Coronal view	0.827 (0.802-0.854)	67.86	73.33	61.29	78.57
	Multiview	**0.933 (0.914-0.952)**	**82.14**	**93.33**	**88.46**	**89.36**
Inception-v3	Axial view	0.910 (0.888-0.933)	82.14	91.11	85.19	89.13
	Sagittal view	0.946 (0.932-0.961)	78.57	91.11	84.61	87.23
	Coronal view	0.921 (0.905-0.937)	85.71	77.78	70.59	89.74
	Multiview	**0.949 (0.945-0.953)**	**82.14**	**95.56**	**92.00**	**89.58**

*Multiview is the combination of axial, sagittal, and coronal planes. The bolded portion is the group with the best indicator.
AUC, area under the curve; CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value.

performance of the classification module based on three single views and the combined views. Among the single-view models, the diagnostic performance on the coronal view was the lowest, whereas it was the best on the axial view, which is inconsistent with previous perceptions (34). These authors (34) suggested that the ABVS-specific coronal view maximizes the understanding of the relationship between the breast lesion and the surrounding tissues and is more conducive to identifying benign and malignant lesions. In particular, the retraction phenomenon on the coronal view has a high sensitivity (80%~89%) and specificity (96%~100%) for detecting breast cancer (11, 35). The main reasons for this contradiction may be that the classification module in this study was constructed on the basis of automatic segmentation, and the relatively poorer segmentation results on the coronal plane led to a subsequent decrease in classification performance. This finding is consistent with a recent view (36) emphasizing that accurate segmentation is a prerequisite for precise classification in DL models. This, in turn, explains the better classification performance of the axial sections. The multiview models simultaneously fused the features of the three views and demonstrated the best diagnostic performance.

Since different CNN backbone network structures may affect the classification performance of the model (37), three common backbone structures (Inception-v3, ResNet50, and MobileNet) were

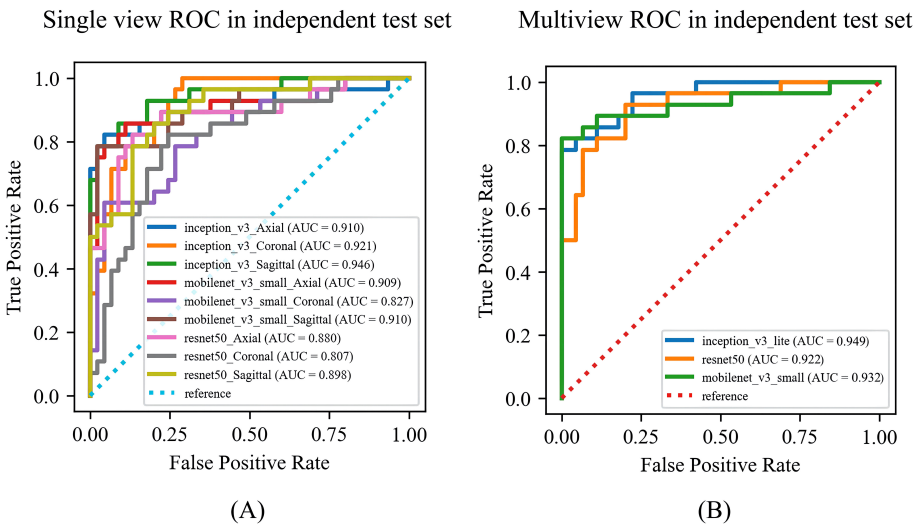


FIGURE 5
The ROC curves of the **(A)** single-view models and **(B)** multiview model based on different backbone networks (ResNet50, MobileNet, and Inception-v3) on the test set. ROC, receiver operating characteristic; AUC, area under the receiver operating characteristic curve.

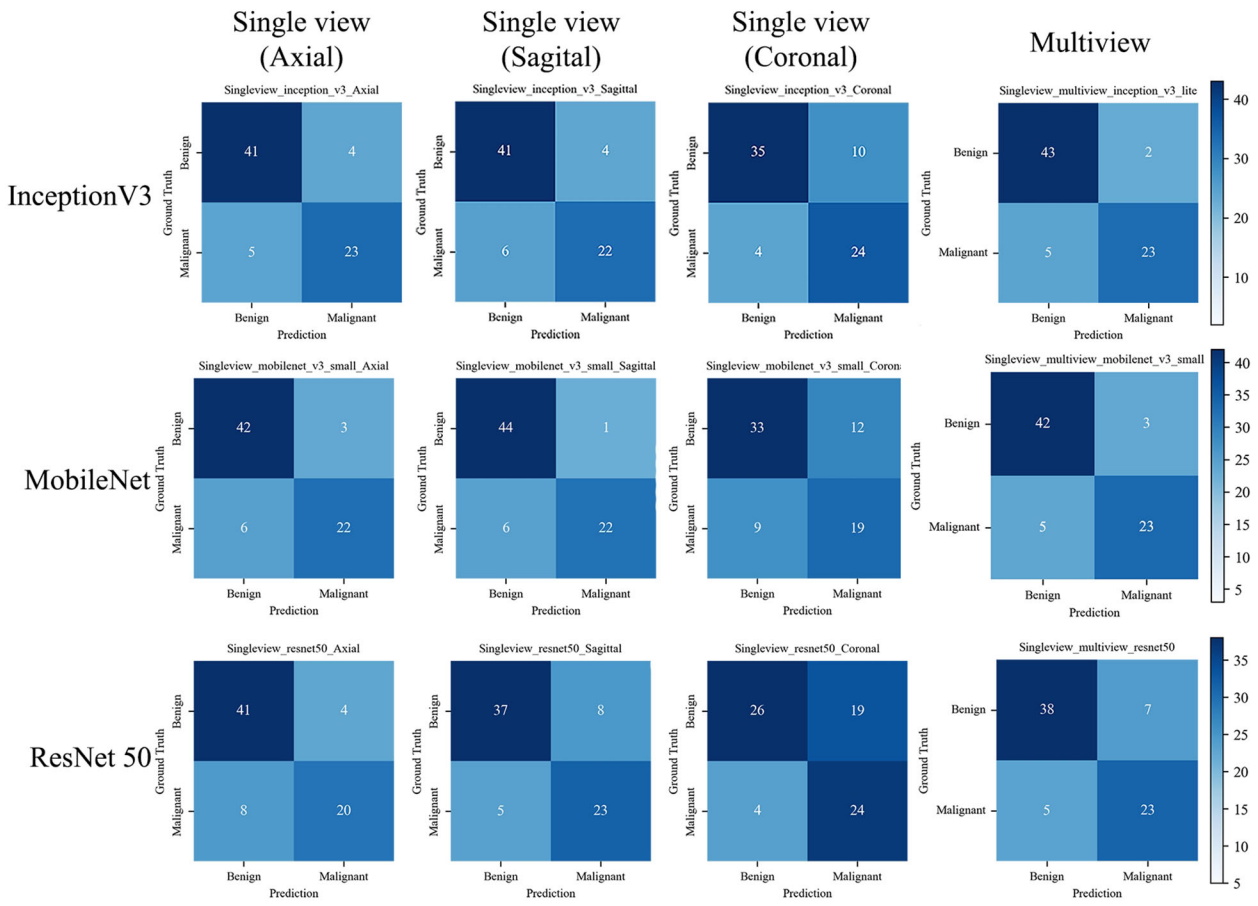


FIGURE 6
The confusion matrix of the deep learning models with the test set. The confusion matrix of the models is based on single view (axial, sagittal, and coronal) and Multiview images, as well as different backbone networks (ResNet50, MobileNet, and Inception v3) with the test set. The correct predictions are shown on the diagonal from the top left to the bottom right of each matrix.

compared for our DL classification modelling. Inception-v3, a CNN improved by the third-generation GoogLeNet, uses multiple regular convolutional layers for feature extraction and concatenates the features as an output, which can help the model learn different-sized lesions efficiently (38). The Inception-v3-based model can handle more and richer spatial features, increase feature diversity, and reduce the number of computations, with an error rate of only 3.5% (39). It actually achieved the highest classification accuracy in the single- and multiview automatic classification models in this study, which may reduce the unnecessary biopsy rate. The optimal DL model in this study could reduce unnecessary biopsies by 53.64%, significantly outperforming previous approaches using contrast-enhanced US (40) and elastography (41), which also have more complicated procedures and rely on the experience of the examining sonographer (42). However, the DL model still had a missed diagnosis rate of 17.85%, which would delay treatment and affect the outcomes and prognoses of patients (43). Although the sagittal single-view model on the MobileNet backbone could reduce the rate of unnecessary biopsies more (57.29%), it was more likely to miss diagnosis (21.43%) than the optimal model was. Therefore, the Inception-v3-based multiview DL model was selected as the final model. On this basis, we further set a decision point with the sensitivity of 100% in the ROC curve which could reduce unnecessary biopsy rate by 21.22% without missing any lesions. However, more comprehensive research and optimization are needed before its application in the clinic. Additionally, previous studies (44, 45) have shown that both the breast lesion itself and its periphery contribute significantly to the interpretability of breast lesions, which is consistent with our Grad-CAM visualization results. To some extent, this may explain the discrimination ability of this DL model.

In conclusion, our work indicated that the ABVS-based DL model can reduce radiologists' manual intervention through automatic segmentation and automatic classification and improve the performance of benign and malignant discrimination of BI-RADS 4 lesions. With further improvements in the model in the future, it will hopefully be promoted and applied in clinical practice, which could significantly impact the management of BI-RADS 4 lesions, reduce biopsies, and promote the development of precision medicine.

There are several limitations. First, the total number of cases of BI-RADS 4 lesions was relatively limited, which may have affected the reliability of the model. Datasets with more centers and larger samples need to be included for further validation and optimization.

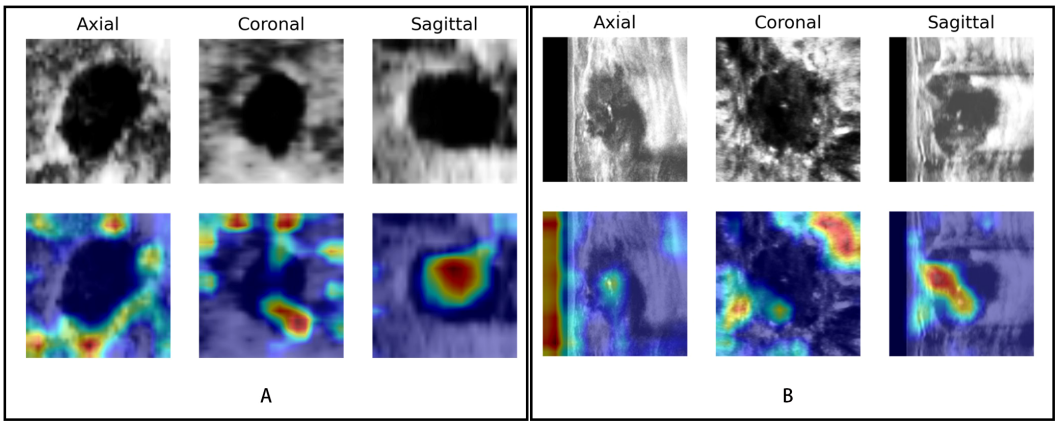


FIGURE 7
Visualization of the DL model. The upper row shows the lesions in different views (axial, coronal, and sagittal) of ABVS images; the row below shows the saliency map generated by the Grad-CAM algorithm. The red region represents a more significant weight. **(A)** A case of a benign BI-RADS 4B lesion and **(B)** a case of a malignant BI-RADS 4C lesion both showed that the DL model focused not only on the lesion itself but also on the periphery of the lesion.

Second, only the largest section of each BI-RADS 4 lesion was used for analysis, which did not fully utilize the advantages of ABVS 3D imaging. The overall information of the lesions is also potentially valuable for predicting their benignity and malignancy. Therefore, the volume of interest in the lesions will be analyzed in a later study. Third, only a single automatic segmentation method from the relevant literature was used in this study. Subsequent studies will explore different automatic segmentation methods to increase the accuracy of model segmentation and further improve model performance.

5 Conclusion

The developed DL model can achieve automatic segmentation and automatic classification of BI-RADS 4 lesions in multiview ABVS images with satisfactory performance. This DL model could reduce the number of unnecessary biopsies of BI-RADS 4 lesions without missing any malignant lesions and simplify the workflow for differential diagnosis, indicating its significant potential for clinical applications.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Institutional Review Board of the Affiliated Hospital of Southwest Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

YL: Conceptualization, Formal analysis, Methodology, Resources, Writing – original draft, Writing – review & editing. CL: Investigation, Methodology, Project administration, Supervision, Writing – review & editing. TY: Investigation, Methodology, Resources, Validation, Writing – review & editing. LC: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing. MH: Investigation, Methodology, Resources, Writing – review & editing. LY: Formal analysis, Software, Writing - review & editing. SZ: Formal analysis, Software, Writing - review & editing. HL: Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – review & editing. JX: Data curation, Methodology, Project administration, Writing – review & editing. SW: Funding acquisition, Investigation, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Sichuan Science and Technology (No.2022YSFS0616) and Doctoral Research Initiation Fund of Affiliated Hospital of Southwest Medical University (No.22161).

Acknowledgments

The authors would like to acknowledge all the study participants and individuals who contributed to the study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1399296/full#supplementary-material

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. (2024) 74:229–63. doi: 10.3322/caac.21834

2. D'Orsi C, Sickles E, Mendelson E, Morris EA, Böhm-Vélez M, Comstock CE, et al. *ACR BI-RADS® Atlas, breast imaging reporting and data system*. American College of Radiology (2013). Available at: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads /Permissions.

3. Elezaby M, Li G, Bhargavan-Chatfield M, Burnside ES, DeMartini WB. ACR BI-RADS assessment category 4 subdivisions in diagnostic mammography: utilization and outcomes in the national mammography database. *Radiology*. (2018) 287:416–22. doi: 10.1148/radiol.2017170770

4. Kerlikowske K, Hubbard RA, Miglioretti DL, Geller BM, Yankaskas BC, Lehman CD, et al. Comparative effectiveness of digital versus film-screen mammography in community practice in the United States: a cohort study. *Ann Intern Med*. (2011) 155:493–502. doi: 10.7326/0003-4819-155-8-201110180-00005

5. Bent CK, Bassett LW, D'Orsi CJ, Sayre JW. The positive predictive value of BI-RADS microcalcification descriptors and final assessment categories. *AJR Am J Roentgenol*. (2010) 194:1378–83. doi: 10.2214/AJR.09.3423

6. Yoon JH, Kim MJ, Moon HJ, Kwak JY, Kim EK. Subcategorization of ultrasonographic BI-RADS category 4: positive predictive value and clinical factors affecting it. *Ultrasound Med Biol*. (2011) 37:693–9. doi: 10.1016/j.ultrasmedbio.2011.02.009

7. Bick U, Trimboli RM, Athanasiou A, Balleyguier C, Baltzer PAT, Bernathova M, et al. Image-guided breast biopsy and localisation: recommendations for information to women and referring physicians by the European Society of Breast Imaging. *Insights Imaging*. (2020) 11:12. doi: 10.1186/s13244-019-0803-x

8. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA*. (2008) 299:2151–63. doi: 10.1001/jama.299.18.2151

9. Shen S, Zhou Y, Xu Y, Zhang B, Duan X, Huang R, et al. A multi-centre randomised trial comparing ultrasound vs mammography for screening breast cancer in high-risk Chinese women. *Br J Cancer*. (2015) 112:998–1004. doi: 10.1038/bjc.2015.33

10. Choi WJ, Cha JH, Kim HH, Shin HJ, Kim H, Chae EY, et al. Comparison of automated breast volume scanning and hand- held ultrasound in the detection of breast cancer: an analysis of 5,566 patient evaluations. *Asian Pac J Cancer Prev*. (2014) 15:9101–5. doi: 10.7314/apjcp.2014.15.21.9101

11. Zheng FY, Yan LX, Huang BJ, Xia HS, Wang X, Lu Q, et al. Comparison of retraction phenomenon and BI-RADS-US descriptors in differentiating benign and Malignant breast masses using an automated breast volume scanner. *Eur J Radiol*. (2015) 84:2123–9. doi: 10.1016/j.ejrad.2015.07.028

12. Wang ZL, Xu JH, Li JL, Huang Y, Tang J. Comparison of automated breast volume scanning to hand-held ultrasound and mammography. *La Radiologia Med*. (2012) 117:1287–93. doi: 10.1007/s11547-012-0836-4

13. Napel S, Mu W, Jardim-Perassi BV, Aerts HJWL, Gillies RJ. Quantitative imaging of cancer in the postgenomic era: Radio(geno)mics, deep learning, and habitats. *Cancer*. (2018) 124:4633–49. doi: 10.1002/cncr.31630

14. Liu H, Chen Y, Zhang Y, Wang L, Luo R, Wu H, et al. A deep learning model integrating mammography and clinical factors facilitates the Malignancy prediction of BI-RADS 4 microcalcifications in breast cancer screening. *Eur Radiol*. (2021) 31:5902–12. doi: 10.1007/s00330-020-07659-y

15. Lin A, Manral N, McElhinney P, Killekar A, Matsumoto H, Kwiecinski J, et al. Deep learning-enabled coronary CT angiography for plaque and stenosis quantification and cardiac risk prediction: an international multicentre study. *Lancet Digit Health*. (2022) 4:e256–65. doi: 10.1016/S2589-7500(22)00022-X

16. Liu F, Liu D, Wang K, Xie X, Su L, Kuang M, et al. Deep learning radiomics based on contrast-enhanced ultrasound might optimize curative treatments for very-early or early-stage hepatocellular carcinoma patients. *Liver Cancer*. (2020) 9:397–413. doi: 10.1159/000505694

17. Kim DY, Choi KH, Kim JH, Hong J, Choi SM, Park MS, et al. Deep learning-based personalised outcome prediction after acute ischaemic stroke. *J Neurol Neurosurg Psychiatry*. (2023) 94:369–78. doi: 10.1136/jnnp-2022-330230

18. Wang Y, Choi EJ, Choi Y, Zhang H, Jin GY, Ko SB. Breast cancer classification in automated breast ultrasound using multiview convolutional neural network with transfer learning. *Ultrasound Med Biol*. (2020) 46:1119–32. doi: 10.1016/j.ultrasmedbio.2020.01.001

19. Cao X, Chen H, Li Y, Peng Y, Wang S, Cheng L. Dilated densely connected U-Net with uncertainty focus loss for 3D ABUS mass segmentation. *Comput Methods Programs BioMed*. (2021) 209:106313. doi: 10.1016/j.cmpb.2021.106313

20. Zhuang Z, Ding W, Zhuang S, Joseph Raj AN, Wang J, Zhou W, et al. Tumor classification in automated breast ultrasound (ABUS) based on a modified extracting feature network. *Comput Med Imaging Graph*. (2021) 90:101925. doi: 10.1016/j.compmedimag.2021.101925

21. *Breast imaging reporting & Data system* . Available online at: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads (Accessed 16 Oct 2023).

22. Kim SH, Kim HH, Moon WK. Automated breast ultrasound screening for dense breasts. *Korean J Radiol*. (2020) 21:15–24. doi: 10.3348/kjr.2019.0176

23. Tan PH, Ellis I, Allison K, Brogi E, Fox SB, Lakhani S, et al. The 2019 World Health Organization classification of tumours of the breast. *Histopathology* (2020) 77 (2):181–5. doi: 10.1111/his.14091

24. Tang JL, Wang Y, Huang CR, Liu H, Al-Nabhan N. Image edge detection based on singular value feature vector and gradient operator. *Math Biosci Eng*. (2020) 17:3721–35. doi: 10.3934/mbe.2020209

25. Xie J, Liu R, Luttrell J, Zhang C. Deep learning based analysis of histopathological images of breast cancer. *Front Genet*. (2019) 10:80. doi: 10.3389/fgene.2019.00080

26. Misra S, Yoon C, Kim KJ, Managuli R, Barr RG, Baek J, et al. Deep learning-based multimodal fusion network for segmentation and classification of breast cancers using B-mode and elastography ultrasound images. *Bioeng Transl Med*. (2023) 8: e10480. doi: 10.1002/btm2.10480

27. Zhang X, Liang M, Yang Z, Zheng C, Wu J, Ou B, et al. Deep learning-based radiomics of B-mode ultrasonography and shear-wave elastography: improved performance in breast mass classification. *Front Oncol*. (2020) 10:1621. doi: 10.3389/fonc.2020.01621

28. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun*. (2020) 11:1236. doi: 10.1038/s41467-020-15027-z

29. Guo X, Liu Z, Sun C, Zhang L, Wang Y, Li Z, et al. Deep learning radiomics of ultrasonography: Identifying the risk of axillary non-sentinel lymph node involvement in primary breast cancer. *EBioMedicine*. (2020) 60:103018. doi: 10.1016/j.ebiom.2020.103018

30. Vakanski A, Xian M, Freer PE. Attention-enriched deep learning model for breast tumor segmentation in ultrasound images. *Ultrasound Med Biol*. (2020) 46:2819–33. doi: 10.1016/j.ultrasmedbio.2020.06.015

31. Lin Y-C, Lin C-H, Lu H-Y, Lin YC, Lin CH, Lu HY, Chiang HJ, Wang HK, Huang YT, et al. Deep learning for fully automated tumor segmentation and extraction of magnetic resonance radiomics features in cervical cancer. *Eur Radiol*. (2020) 30:1297–305. doi: 10.1007/s00330-019-06467-3

32. Wu S, Li H, Quang D, Guan Y. Three-plane-assembled deep learning segmentation of gliomas. *Radiol Artif Intell*. (2020) 2:e190011. doi: 10.1148/ryai.2020190011

33. Avanzo M, Wei L, Stancanello J, Vallières M, Rao A, Morin O, et al. Machine and deep learning methods for radiomics. *Med Phys*. (2020) 47:e185–202. doi: 10.1002/mp.13678

34. Girometti R, Zanotel M, Londero V, Linda A, Lorenzon M, Zuiani C. Automated breast volume scanner (ABVS) in assessing breast cancer size: A comparison with conventional ultrasound and magnetic resonance imaging. *Eur Radiol*. (2018) 28:1000–8. doi: 10.1007/s00330-017-5074-7

35. van Zelst JCM, Mann RM. Automated three-dimensional breast US for screening: technique, artifacts, and lesion characterization. *Radiographics*. (2018) 38:663–83. doi: 10.1148/rg.2018170162

36. Xie Y, Zhang J, Lu H, Shen C, Xia Y. SESV: accurate medical image segmentation by predicting and correcting errors. *IEEE Trans Med Imaging*. (2021) 40:286–96. doi: 10.1109/TMI.2020.3025308

37. Chen S, Morales-Sanfrutos J, Angelini A, Cutting B, Heinis C. Structurally diverse cyclisation linkers impose different backbone conformations in bicyclic peptides. *Chembiochem*. (2012) 13:1032–8. doi: 10.1002/cbic.201200049

38. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vis*. (2015) 115:211–52. doi: 10.1007/s11263-015-0816-y

39. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2016) p. 2818–26.

40. Park S-Y, Kang BJ. Combination of shear-wave elastography with ultrasonography for detection of breast cancer and reduction of unnecessary biopsies: a systematic review and meta-analysis. *Ultrasonography*. (2021) 40:318–32. doi: 10.14366/usg.20058

41. Golatta M, Pfob A, Büsch C, Bruckner T, Alwafai Z, Balleyguier C, et al. The potential of combined shear wave and strain elastography to reduce unnecessary biopsies in breast cancer diagnostics - An international, multicentre trial. *Eur J Cancer*. (2022) 161:1–9. doi: 10.1016/j.ejca.2021.11.005

42. Park S-Y, Kang BJ. Combination of shear-wave elastography with ultrasonography for detection of breast cancer and reduction of unnecessary biopsies: a systematic review and meta-analysis. *Ultrasonography*. (2021) 40:318–32. doi: 10.14366/usg.20058

43. Bae MS, Kim H-G. Breast cancer risk prediction using deep learning. *Radiology*. (2021) 301:559–60. doi: 10.1148/radiol.2021211446

44. Xu H, Liu J, Chen Z, Wang C, Liu Y, Wang M, et al. Intratumoral and peritumoral radiomics based on dynamic contrast-enhanced MRI for preoperative prediction of intraductal component in invasive breast cancer. *Eur Radiol*. (2022) 32:4845–56. doi: 10.1007/s00330-022-08539-3

45. Gu Y, Xu W, Liu T, An X, Tian J, Ran H, et al. Ultrasound-based deep learning in the establishment of a breast lesion risk stratification system: a multicenter study. *Eur Radiol*. (2023) 33:2954–64. doi: 10.1007/s00330-022-09263-8

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 20 December 2024
DOI 10.3389/fonc.2024.1437506

OPEN ACCESS

EDITED BY
Jiabao He,
Newcastle University, United Kingdom

REVIEWED BY
Bernard Siow,
The Francis Crick Institute, United Kingdom
Simon Lowes,
Gateshead Health NHS Foundation Trust, United Kingdom

*CORRESPONDENCE
Savannah C. Partridge
scp3@uw.edu

RECEIVED 23 May 2024
ACCEPTED 02 December 2024
PUBLISHED 20 December 2024

CITATION
Biswas D, Hippe DS, Winter AM, Li I, Rahbar H and Partridge SC (2024) Diffusion weighted imaging for improving the diagnostic performance of screening breast MRI: impact of apparent diffusion coefficient quantitation methods and cutoffs. *Front. Oncol.* 14:1437506. doi: 10.3389/fonc.2024.1437506

Diffusion weighted imaging for improving the diagnostic performance of screening breast MRI: impact of apparent diffusion coefficient quantitation methods and cutoffs

Debosmita Biswas[1,2], Daniel S. Hippe[3], Andrea M. Winter[1], Isabella Li[1], Habib Rahbar[1] and Savannah C. Partridge[1,2]*

[1]Department of Radiology, School of Medicine, University of Washington, Seattle, WA, United States, [2]Department of Bioengineering, College of Engineering, University of Washington, Seattle, WA, United States, [3]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

Introduction: Diffusion weighted MRI (DWI) has emerged as a promising adjunct to reduce unnecessary biopsies prompted by breast MRI through use of apparent diffusion coefficient (ADC) measures. The purpose of this study was to investigate the effects of different lesion ADC measurement approaches and ADC cutoffs on the diagnostic performance of breast DWI in a high-risk MRI screening cohort to identify the optimal approach for clinical incorporation.

Methods: Consecutive screening breast MRI examinations (August 2014–Dec 2018) that prompted a biopsy for a suspicious breast lesion (BI-RADS 4 or 5) were retrospectively evaluated. On DWI, ADC (b=0/100/600/800s/mm^2) measures were calculated with three different techniques for defining lesion region-of-interest (ROI; single slice('2D'), whole volume('3D') and lowest ADC region ('hotspot')). An optimal data-derived ADC cutoff for each technique was retrospectively identified to reduce benign biopsies while avoiding any false negatives, inherently producing cutoffs with 100% sensitivity in this particular cohort. Further, diagnostic performance of these measures was validated using two prespecified ADC cutoffs: 1.53x10^{-3}mm^2/s from the ECOG-ACRIN A6702 trial and 1.30x10^{-3}mm^2/s from the international EUSOBI group. Diagnostic performance was compared between ADC maps generated with 2(0/800s/mm^2) and 4(0/100/600/800s/mm^2) b-values. Benign biopsy reduction rate was calculated (number of benign lesions with ADC >cutoff)/(total number of benign lesions).

Results: 137 suspicious lesions (in 121 women, median age 44 years [range, 20-75yrs]) were detected on contrast-enhanced screening breast MRI and recommended for biopsy. Of those, 30(21.9%) were malignant and 107(78.1%) were benign. Hotspot ADC measures were significantly lower (p<0.001) than ADCs from both 2D and 3D ROI techniques. Applying the optimal data-derived ADC cutoffs resulted in comparable reduction in benign biopsies across ROI techniques (range:16.8% -17.8%). Applying the prespecified A6702 and EUSOBI cutoffs resulted in benign biopsy reduction rates of 11.2-19.6%(with 90.0-100%

sensitivity) and 36.4-51.4%(with 70.0-83.3% sensitivity), respectively, across ROI techniques. ADC measures and benign biopsy reduction rates were similar when calculated with only 2 b-values (0,800 s/mm^2) versus all 4 b-values.

Discussion: Our findings demonstrate that with appropriate ADC thresholds, comparable reduction in benign biopsies can be achieved using lesion ADC measurements computed from a variety of approaches. Choice of ADC cutoff depends on ROI approach and preferred performance tradeoffs (biopsy reduction vs sensitivity).

KEYWORDS

diffusion weighted imaging (DWI), apparent diffusion coefficient (ADC), breast magnetic resonance imaging (MRI), diagnostic performance, false positives, ADC cutoff, region-of-interest (ROI)

Introduction

Breast cancer is the most common type of cancer and second leading cause of cancer deaths in women in the United States (1). Timely detection of cancer can lead to better treatment outcomes and higher survival rates for patients, making breast cancer screening a crucial aspect of women's health. It is well-established that breast MRI offers superior sensitivity for detecting breast cancer versus other clinical breast imaging techniques and is therefore recommended for screening of high risk women (2–4). The high sensitivity of breast MRI relies on injection of intravenous contrast to identify areas of suspicious vasculature, commonly associated with breast malignancies. Dynamic contrast enhanced breast MRI (DCE) provides high sensitivity (> 85%) for breast cancer detection but suffers from moderate specificity, resulting in unnecessary biopsies that cause needless expense, inconvenience, discomfort and emotional distress for the patient (5, 6). Diffusion weighted imaging (DWI) is a non-contrast functional MRI technique that provides information based on microscopic movement of water molecules in tissues and allows an indirect assessment of tissue microstructure and cellularity. Breast malignancies tend to restrict diffusion and DWI has shown clear potential to increase breast MRI diagnostic specificity when used along with DCE.

The apparent diffusion coefficient (ADC), derived from DWI, is commonly used to quantify *in vivo* diffusion. Numerous studies have reported the utility of the metric for distinguishing between benign and malignant breast findings, suggesting ADC cutoff values could be safely used to downgrade suspicious enhancing lesions and avoid unnecessary biopsies (7–12). However, diffusion-weighted MRI is not yet incorporated into the Breast Imaging Reporting and Data System (BI-RADS) (13), and more data are needed to refine optimal methods for clinical implementation, particularly regarding quantitation of lesion ADC. Approaches to measure lesion ADC values vary and emphasize different aspects of the tumor microstructure. Choice of region-of-interest (ROI) sampling methods capture different aspects of the lesion (e.g., whole volume of the lesion to comprehensively measure the entire tumor versus 'hotspot' for peak cellularity) (9, 14). ADC is most commonly reported as the mean value across the lesion, measured using a manually defined ROI from a single slice. Alternative approaches of obtaining lesion ADC values include utilizing multiple (more than 2) b-values to compute the ADC map, using a nonzero minimum b-value to reduce confounding perfusion effects in ADC calculation, segmenting the whole 3D volume of the tumor across multiple slices to better account for cellular and microstructural heterogeneity across the abnormality (15), and measuring just the subregion of greatest diffusion restriction within the tumor potentially reflecting highest cellularity and proliferation within the tumor (16).

Despite evidence of ADC as a valuable biomarker for diagnosing breast cancer from multicenter prospective (10) and retrospective studies (17), implementation of DWI into routine clinical interpretations is still a work in progress. Lack of standardization of acquisition protocols and variability in ROI definition techniques and study population have resulted in a broad range of reported ADC values and diagnostic thresholds, hindering clinical integration of DWI as a screening tool. Most of the prior studies evaluated data derived from patients who received breast MRI for diagnostic purposes (to evaluate symptomatic breast tumors, abnormalities detected on other imaging modalities, or to evaluate extent of disease for known cancers) rather than for asymptomatic screening. Lesions detected in screening breast MRI exams are usually smaller and may not exhibit the same characteristics as symptomatic breast tumors, which have been used to determine the ADC cutoffs in many prior studies.

Therefore, this study aimed to investigate how different methods of measuring lesion ADC values affect the diagnostic performance of breast MRI in a high-risk screening cohort. To our knowledge, no prior research has focused exclusively on lesions

identified through screening MRI. This distinctive cohort allows us to identify the most effective measurement approach for clinical incorporation of DWI in breast screening.

Materials and methods

Participants

The institutional review board approved this single academic medical center retrospective study (Fred Hutchinson Cancer Center institutional review no. 7339). Requirement for informed consent was waived for reviewing clinical images and medical records. Consecutive screening breast MRIs between May 2015 and December 2018 with a biopsy recommendation (BI-RADS 4 or 5 assessment) followed by definitive biopsy outcome were included in this study. Medical records were reviewed to determine two year follow-up for lesions with benign pathology on biopsy. All breast MRIs were prospectively interpreted by one of several fellowship-trained breast radiologists (including HR, with over 10 years of breast imaging experience). Over this timeframe, DWI was not generally used for BI-RADS assessment due to a lack of consensus on an ADC threshold to obviate biopsy. Lesion outcomes were classified as benign or malignant based on pathology reports after breast biopsy or excision. A subset of participants in our study (n = 108) were previously described in another study validating the diagnostic performance of point of care (recorded in the clinic) ADC measures of MRI detected breast lesions using pre-specified cutoffs (11). In this study, we evaluated the effects of different b-values and ROI segmentation techniques on ADC performance for reducing unnecessary biopsies in breast screening exams.

MRI acquisition

All breast MRI examinations were acquired on a 3T MR scanner (Achieva Tx; Philips Healthcare, Best, Netherlands) with a dedicated 16-channel breast coil. Images were acquired in the axial orientation, and each exam included T2-weighted, DWI, and DCE-MRI sequences, in accordance with American College of Radiology (ACR) breast MRI accreditation and European Society of Breast Imaging (EUSOBI) breast DWI guidelines (18), and following the ECOG-ACRIN A6702 DWI protocol (10) (full protocol in Table 1). Onboard software provided by the scanner manufacturer was used for both spatially registering the DW images across b-values to correct for patient motion and eddy current effects.

Image analysis

For the primary analysis, ADC maps were computed using a classic monoexponential decay model and least squares fitting of the signal decay across all b-values up to b = 800 s/mm^2 (b =0, 100, 600, 800 s/mm^2) as recommended by EUSOBI consensus for standardized reporting of breast ADC values (18, 19). The following equation was used

$$S_b = S_0 * e^{-b*ADC}$$

where S_b is the DWI intensity signal at weighting b, S_0 is the signal intensity with no diffusion weighting and ADC expressed in mm^2/s. ADC maps were also computed using only 2 b-values (b=0, 800 s/mm^2) for secondary analysis. Lesion ROIs were defined by a researcher (DB) guided by a fellowship trained radiologist (AW) who did not participate in the prospective reads, all blinded to biopsy outcomes. Using a semi-automated threshold-based software tool to avoid fat and fibroglandular tissue developed in MATLAB (Mathworks, Natick, MA) (20), lesion ROIs were defined for a single representative slice ('2D') and whole tumor volume across multiple slices ('3D') on the b = 800 s/mm^2 images and then propagated to ADC maps. For lesions measurable only on a single slice, the 2D and 3D measurements will be the same. A subregion (9 -16 contiguous pixels, depending on lesion morphology) within the 3D lesion ROI producing the lowest mean ADC value was automatically selected by the software as the 'hotspot', following consensus recommendations (18). For each lesion ROI, the mean ADC of all voxels was calculated for primary analysis, while other histogram metrics (minimum, maximum, standard deviation, etc) were also calculated for 3D ROIs.

ADC thresholds

For optimal clinical integration and patient safety, our study focused on ADC cutoffs that could reduce false positives while still maintaining high sensitivity (minimizing false negatives). ADC cutoffs were retrospectively determined for each of the ADC measurement techniques based on the highest ADC observed among malignant lesions, as previously described (10). For each ADC technique, lesions with ADC measures above the cutoff would be considered probably benign and avoid biopsy. These optimal data-derived ADC cutoffs will inherently achieve 100% sensitivity in the current dataset because they were selected retrospectively, though 100% sensitivity may not be achieved with the same cutoffs in another cohort. The primary purpose of selecting these ADC cutoffs was to enable comparison of biopsy reduction rates of the three ROI techniques at a comparable and clinically relevant operating point, where sensitivity is held fixed at 100%. We further evaluated the performance of two previously proposed ADC cutoffs: 1) 1.53 x 10^{-3} mm^2/s determined by the ECOG ACRIN A6702 multicenter study (10) and 2) 1.30 x 10^{-3} mm^2/s recommended by EUSOBI consensus guidelines (18, 21) and implemented in a large prospective DWI screening trial (22).

Statistical analysis

The analysis was performed at the lesion level. Paired comparison of mean ADC values between ROI techniques or ADC maps (based on 4 b-values vs. 2 b-values) were performed using generalized estimating equations (GEE) based regression to account for non-independence of multiple lesions from the same patient (23). ADC values were also compared between benign and malignant lesions using GEE-based

10.3389/fonc.2024.1437506

TABLE 1 MRI protocol parameters.

	DWI	T2-weighted	DCE
Sequence type	Diffusion-weighted spin echo, echo planar imaging (DW SE-EPI)	Turbo spin echo (TSE)	Fast Field Echo (FFE)
2D or 3D sequence	2D	2D	3D
Slice orientation	Axial	Axial	Axial
Laterality	Bilateral	Bilateral	Bilateral
Phase direction	A/P	R/L	R/L
FOV	360 mm x 360 mm	240 mm x 360 mm	360 mm x 360 mm
In-plane Resolution	1.8 mm x 1.8 mm	1 mm x 1 mm	0.5 mm x 0.5 mm
Slice thickness	4 mm	3 mm	1.3 mm
Fat-suppression	SPAIR	SPAIR	SPAIR
TR	5000 ms	5000 ms	5.95 ms
TE	60 ms	60 ms	3 ms
Echo Train Length	67	N/A	N/A
Flip Angle	90 degrees	90 degrees	10 degrees
b-values	0, 100, 600, 800, 1000 s/mm^2	N/A	N/A
Number of slices	30	~60; Variable; complete bilateral coverage	~150/ Variable; complete bilateral coverage
Slice Gap	No gap	No gap	No gap
Parallel imaging factor	3 Phase	3.1 Phase	2.7 Phase, 2 Slice
No. of averages	2 (b=0, 100), 4 (b=600, 800), 6 (b=1000)	1	1
Contrast injection	N/A	N/A	Intravenous injection of 0.1mmol/kg body-weight gadoteridol
Sequence acquisition time	4:30 minutes	2:45 minutes	2:54 mins per phase, 12 mins total (1 pre, 3 post-contrast phases with k0 at ~2, 5, and 8 minutes after contrast injection)
Diffusion Gradient Parameters			
Amplitude Duration Separation #Directions	22.52 mT/m 13 ms 21 ms 3	N/A	N/A

SPAIR, Spectral attenuated inversion recovery. N/A, Not Applicable.

regression. Diagnostic performance of each ROI measure was summarized using the area under the receiver operating characteristic curve (AUC), sensitivity (proportion of malignant lesions with ADC ≤ cutoff), and benign biopsy reduction rate (proportion of benign lesions with ADC > cutoff). Confidence intervals (CIs) were computed using the non-parametric bootstrap (24) or GEE-based regression, clustered by patient, to account for non-independence of multiple lesions from the same patient. CIs for sensitivity were calculated using the Clopper-Pearson exact method (25) due to the smaller sample size of malignant lesions and only two patients had multiple malignant lesions (two lesions each). Benign biopsy reduction rates were compared between ROI techniques using the sign test and between lesion subgroups (by lesion type or lesion size) using GEE-based regression. The performance of the data-driven thresholds was further explored using 5-fold cross-validation, where different random subsamples of patients were used to rederive cutoffs and subsequently test the cutoffs in held-out subsamples not used in selecting the cutoff. The cross-validations were repeated 1,000 times and the results were averaged. Statistical significance was defined as two-sided $p< 0.05$. All analyses were performed using R version 4.0.3.

Results

During the study period, 2329 screening breast MRI examinations were performed and 137 BI-RADS category 4/5 lesions were detected in 116 women (median age, 46 years, range [20-75 years]) who underwent biopsy. Pathologic assessment

revealed 30 malignancies (21.9% of lesions [30/137]; 12 invasive ductal carcinoma, 13 ductal carcinoma *in situ* [DCIS], 4 invasive lobular carcinoma [ILC], and 1 malignant phyllodes tumor) and 107 benign (including 3 high-risk lesions with atypical ductal hyperplasia, lobular carcinoma *in situ*, and atypical lobular hyperplasia) lesions. None of the benign lesions upgraded to malignancy within the two year follow up period. The median size of the lesions was 8 mm (range: 3 – 76 mm), while 85% (20/137) of lesions were only measurable on a single slice due to small size or avoidance of partial volume averaging effects. Sixty-six lesions were non mass enhancements (NME) and 71 were masses (Table 2).

The mean and standard deviation ADC measures for each of the ROI definition techniques were 1.27 ± 0.35, 1.26 ± 0.35, and 1.16 ± 0.36 x 10^{-3} mm^2/s for 2D, 3D, and hotspot, respectively. Pairwise comparisons revealed that hotspot ADC measures were significantly lower than 2D and 3D segmentations (mean ADC difference = 0.09 x 10^{-3} (mm^2/s), respectively, $p<0.001$ for both) while there was only a small, but statistically significant, difference in ADC measures between 2D and 3D segmentations (mean ADC difference = 0.01 x10^{-3} mm^2/s, $p = 0.020$).

Mean ADC measures were significantly higher for benign (range, 1.21 to 1.31 x10^{-3} mm^2/s) versus malignant lesions (0.97 to 1.11 x10^{-3} mm^2/s) by all three ROI techniques ($p< 0.001$ for each, examples, Figures 1, 2). AUCs for predicting malignancy were similar for the three ROI techniques (2D: 0.66 [95% CI: 0.55-0.77], 3D: 0.67 [95% CI: 0.56-0.77], hotspot: 0.68 [95% CI: 0.57-0.79], Table 3). No other histogram metrics measured from the ROIs demonstrated improved performance over mean ADC (Supplementary Table 1).

TABLE 2 Subject and lesion characteristics.

	N (%) or Median (Range)
Women (N total)	116
Lesions (N total)	137
Mean Age (years)	46 (20-75)
Race	
White	100 (86.2%)
Asian	7 (6.0%)
Black	2 (1.7%)
American Indian / Alaskan Native	1 (0.9%)
Native Hawaiian/Other Pacific Islander	1 (0.9%)
Unknown	5 (4.3%)
Ethnicity	
Hispanic/Latino	2 (1.7%)
Not Hispanic/Latino	107 (92.2%)
Unknown	7 (6.0%)
Primary MRI Screening Indication	
Personal History	35 (30.2%)
Genetic Mutation/ Family History	71 (61.2%)
Other (eg, prior atypia diagnosis)	10 (8.6%)
Menopausal Status	
Pre	62 (53.4%)
Post	54 (46.6%)
Lesion size*	
≤10mm	67 (48.9%)
>10mm	70 (51.1%)
BI-RADS Assessment*	
Category 4	135 (98.5%)
Category 5	2 (1.5%)
Lesion type*	
Mass	71 (51.8%)
NME	66 (48.2%)
Method of biopsy*	
MRI guided needle biopsy	100 (73%)
Ultrasound guided biopsy	35 (26%)
Stereotactic biopsy	2 (1%)
Pathology Outcome*	
Malignant	30 (21.9%)
Invasive ductal carcinoma Ductal carcinoma in situ Invasive lobular carcinoma Malignant phyllodes	13 12 4 1
Benign	107 (78.1%)

*Calculated at lesion level (N = 137), otherwise at patient level (N = 116).

Data-derived ADC thresholds and diagnostic performance

The optimal ADC cutoffs derived from the data (producing 100% sensitivity) resulted in the same cutoff value for 3D and 2D ROIs (1.55 x 10^{-3} mm^2/s) while the cutoff for hotspot was lower (1.44 x 10^{-3} mm^2/s) (Figures 3, 4; Table 4). Applying these data-derived ADC cutoffs resulted in a 17.8% (19/107) reduction in benign biopsies using 2D (95% CI: 10.4-25.1%) and hotspot ROIs (95% CI: 10.0-25.5%), and 16.8% (18/107, 95% CI: 9.6-24.0%) for

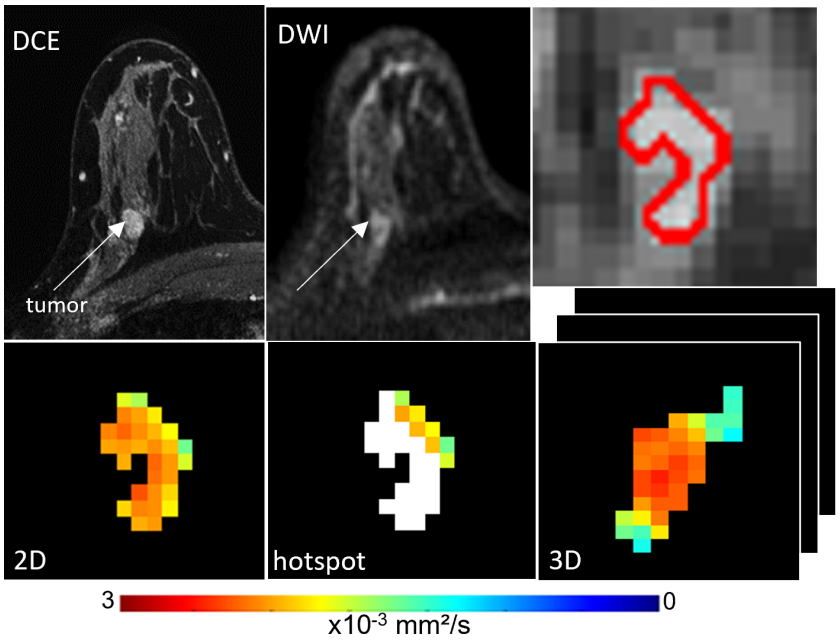


FIGURE 1
ADC measures of a BI-RADS 4 11 mm mass detected in a 41-year-old woman who underwent screening MRI. Lesion ADCs calculated using the different ROI techniques were 2.12, 1.84, and 1.97 x10^{-3} mm^2/s for 2D, hotspot, and 3D ROIs, respectively. On biopsy, it was found that the lesion was benign breast tissue with focal fibrocystic changes.

3D ROIs ($p > 0.99$ for each pairwise comparison between ROI techniques).

Performance of the data-derived cutoffs based on repeated cross-validations was explored in Supplementary Table 2. Average sensitivity estimated from held-out subsamples not used to derive cutoffs ranged from 95.9% to 96.0% across ROI techniques. Average benign biopsy reduction rates across the held-out subsamples were similar to their values based on all of the data, ranging from 17.3% to 19.6% across techniques.

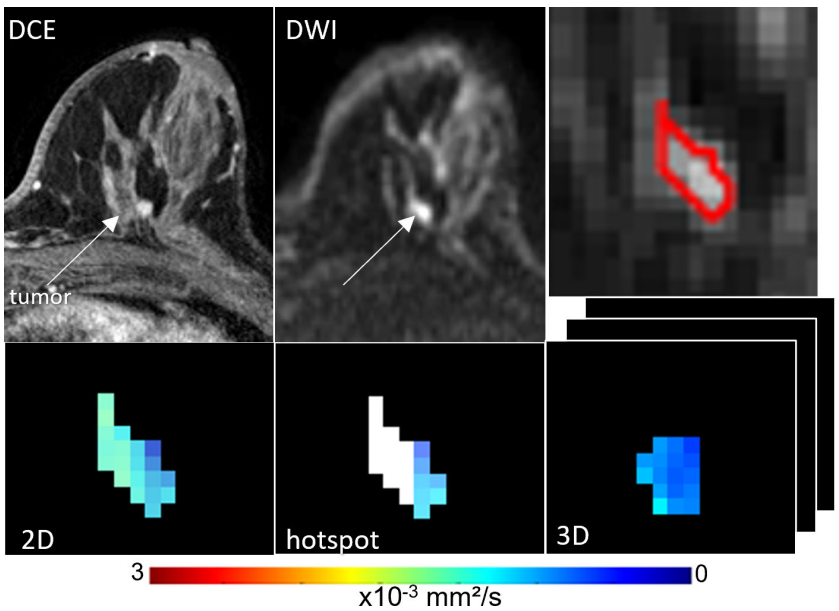


FIGURE 2
ADC measures of a BI-RADS 4 6 mm mass detected in a 55-year-old woman who underwent screening MRI. Lesion ADCs calculated using the different ROI techniques were 1.21, 1.17 and 1.22 x10^{-3} mm^2/s for 2D, hotspot, and 3D ROIs, respectively. The lesion was invasive ductal carcinoma on biopsy.

TABLE 3 ADC for differentiating benign and malignant lesions using different region of Interest (ROI) techniques.

ROI Technique	ADC measures Mean ± SD (x10^{-3} mm^2/s)		P-value	AUC (95% CI)
	Malignant N = 30	Benign N = 107		
2D	1.11 ± 0.29	1.31 ± 0.35	<0.001	0.66 (0.55-0.77)
3D	1.10 ± 0.29	1.31 ± 0.35	<0.001	0.67 (0.56-0.77)
Hotspot	0.97 ± 0.32	1.21 ± 0.35	<0.001	0.68 (0.57-0.79)

ADC, apparent diffusion coefficient; AUC, area under the curve; CI, confidence interval; ROI, region of interest.

Stratified performance of ADC measure

Stratifying by size, no significant differences were observed in benign biopsy reduction rates between larger lesions (19.6% [10/51] to 21.5% [11/51]) and smaller lesions (14.3% [8/56] to 16.1% [9/56]) for each ROI technique ($p > 0.36$ for each) (Table 5). Similarly, there were no significant differences in performance between ROI techniques for either larger lesions ($p > 0.99$ for each pairwise comparison) or smaller lesions ($p > 0.99$).

Stratifying by lesion type, benign biopsy reduction rates from each ROI technique were comparable between NMEs (17.0% [8/47] to 21.2% [10/47]) and masses (15.0% [9/60] to 18.3% [11/60], $p > 0.44$ for each) (Table 5). There were no significant differences in performance using the three ROI approaches by lesion type. For masses: 2D ROI (18.3% [11/60]), hotspot (15.0% [9/60]) and 3D ROI (16.7% [10/60]), ($p > 0.62$ for each pairwise comparison) and for NMEs: 2D and 3D ROI (17.0% [8/47], hotspot (21.2% [10/47], ($p = 0.69$).

Diagnostic performance of prespecified cutoffs

A6702 cutoff (1.53 x 10^{-3} mm^2/s): This cutoff was determined in the A6702 trial to reduce biopsies while prioritizing sensitivity, with lesion ADC values generated using 4 b-values (same b-values as this study) and a 3D ROI approach. In this dataset, applying the A6702 cutoff to ADC values measured using the same approach as A6702 resulted in 19.6% reduction (21/107, 95% CI: 12.1-27.1%) in benign biopsies with 93.3% sensitivity (28/30, 95% CI: 77.9-99.2%). Results were similar for 2D ROI measures, with 19.6% reduction in benign biopsies (21/107, 95% CI: 12.1-27.1%) and 90.0% sensitivity (27/30, 95% CI: 73.5-97.9%). For hotspot ROIs, the A6702 cutoff was notably higher than data derived optimal ADC 100% sensitivity cutoff (1.44 x 10^{-3} mm^2/s) and avoided fewer benign biopsies (11.2% [12/107] vs. 17.8% [17/107]).

EUSOBI cutoff (1.3 x 10^{-3} mm^2/s): The EUSOBI working group recommends using a focused ROI in the area of lowest ADC within the enhancing lesion, similar to the hotspot ROI approach in this study. Applying the EUSOBI cutoff to hotspot ROI ADC measures in this dataset led to a 36.4% reduction in benign biopsies (39/107, 95% CI: 27.2-45.7%) and 83.3% sensitivity (25/30, 95% CI: 65.3-94.4%). However, applying the EUSOBI cutoff to 2D and 3D ROI measures resulted in a very high reduction in benign biopsies (50.5% [54/107] and 51.4% [55/107], respectively) but substantially lowered sensitivity (70.0% [21/30] and 73.3% [22/30]).

Secondary analysis of ADC mapping using two vs four b-values

Pairwise comparisons revealed that ADC measures computed from only 2 b-values (0, 800 s/mm^2) were not significantly different

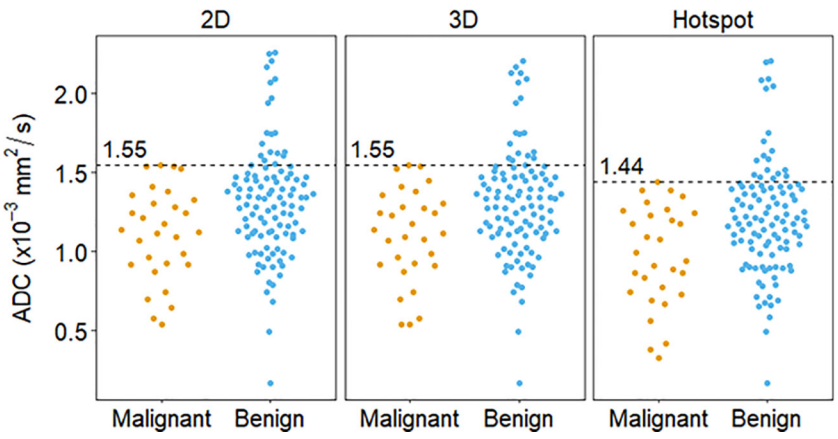


FIGURE 3
Distribution of lesion ADC measures calculated via different ROI techniques and the corresponding cutoff (dashed lines), derived using the highest malignant ADC value (100% sensitivity).

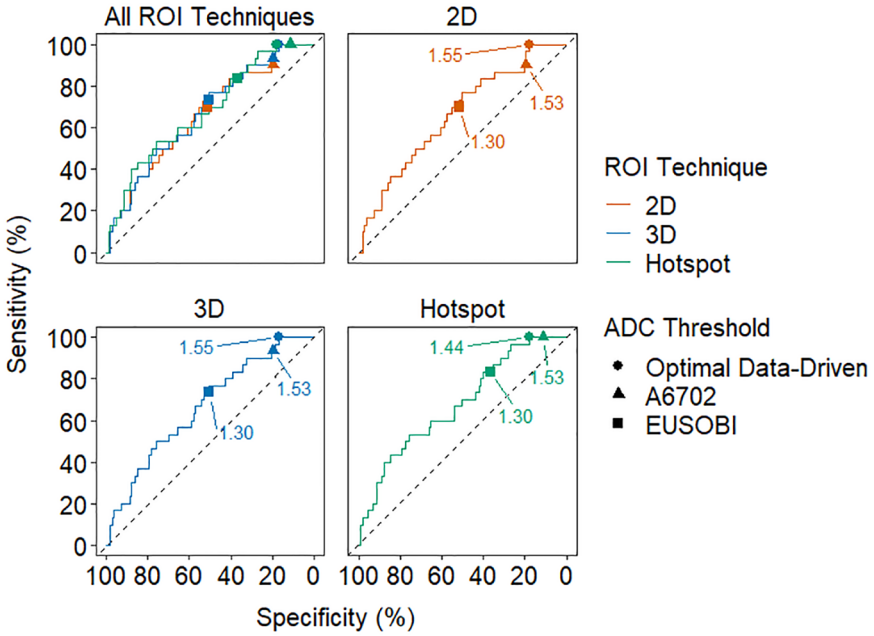


FIGURE 4
Receiver operating characteristic (ROC) curves for each ROI technique. The optimal data-driven threshold for each technique is marked with a circle on each curve and the A6702 and EUSOBI thresholds are marked with a triangle and square, respectively. The thresholds are labeled in units of 10^{-3} mm^2/s.

from measures computed using all four b-values (0, 100, 600, 800 s/mm^2) on average (|mean ADC difference|< 0.01 x10^{-3} mm^2/s for each pairwise comparison between ROI techniques, p > 0.17 for each). Similarly, diagnostic performance of ADC measures computed from 2 b-values were almost identical to that from ADC measured computed from 4 b-values for all ROI techniques (Table 6). For example, the benign biopsy reduction rates from the A6702 cutoff were 19.6% vs. 19.6% (4 b-values vs. 2 b-values) for 2D ROIs and 19.6 vs. 18.7% for 3D ROIs and benign biopsy reduction rates from the EUSOBI recommended cutoff were 36.4% vs 34.6% for hotspot ROIs.

Discussion

Suspicious enhancement of normal parenchymal tissue and benign tumors on DCE MRI leads to benign findings in as many as four in five screening-MRI prompted biopsies. Reducing false positives and unnecessary biopsies is of high importance due to the growing utilization of breast MRI for screening women with elevated breast cancer risk. At the same time, maintaining the high sensitivity of breast MRI is critical to ensure its value for early detection of disease. Although many studies have shown a clear potential of utilizing DWI for improving the diagnostic performance of breast MRI with minimum increase in cost and scan time, it has not yet been incorporated into BI-RADS. One challenge lies in determining the optimal approach for standardized integration of DWI in the clinic. Therefore, our study investigated the effect of various ADC measurement approaches and ADC cutoffs (data derived and prespecified) on performance to reduce unnecessary breast biopsies in lesions detected on breast MRI screening exams. Overall, our results using data-derived cutoffs showed that a variety of ADC measurement techniques could significantly distinguish benign and malignant breast lesions (AUCs 0.66 – 0.68) and reduce the rate of unnecessary biopsies (by 17% to 18%) of conventional breast MRI without missing any cancers. Results using prespecified cutoffs further illustrated the importance of performing lesion ADC measurements consistent to that by which the cutoffs were derived in order achieve maximal performance.

Selection of ADC threshold depends on clinical preferences regarding tradeoffs between sensitivity and specificity. Many prior

TABLE 4 Optimal data-derived ADC cutoffs maintaining 100% sensitivity.

ROI Technique	Optimal ADC Cutoff (x10^{-3} mm^2/s)	Benign Biopsy Reduction Rate		
		No.	Estimate	(95% CI)
2D	1.55	19/107	17.8%	(10.4, 25.1%)
3D	1.55	18/107	16.8%	(9.6, 24.0%)
Hotspot	1.44	19/109	17.8%	(10.0, 25.5%)

ADC, apparent diffusion coefficient; CI, confidence interval; ROI, region of interest.

TABLE 5 Subgroup analysis of ADC performance by lesion type and size.

		Benign Biopsy Reduction Rate		
Subgroup	ROI Technique	No.	Estimate	(95% CI)
Masses	2D	11/60	18.3%	(7.8, 28.9%)
	3D	10/60	16.7%	(6.4, 26.9%)
	Hotspot	9/60	15.0%	(5.3, 24.7%)
NMEs	2D	8/47	17.0%	(6.5, 27.6%)
	3D	8/47	17.0%	(6.5, 27.6%)
	Hotspot	10/47	21.3%	(8.9, 33.6%)
Lesion size ≤ 10 mm	2D	8/56	14.3%	(4.2, 24.4%)
	3D	8/56	14.3%	(4.2, 24.4%)
	Hotspot	9/56	16.1%	(5.6, 26.6%)
Lesion size > 10 mm	2D	11/51	21.6%	(10.4, 32.8%)
	3D	10/51	19.6%	(8.8, 30.4%)
	Hotspot	10/51	19.6%	(9.4, 29.8%)

ADC, apparent diffusion coefficient; CI, confidence interval; NME, non-mass enhancement; ROI, region of interest.

studies have defined an ADC cutoff by equally optimizing sensitivity and specificity (26–28). While a false positive finding can lead to further testing (biopsy) and unnecessary emotional distress to the patient, a false negative is potentially more detrimental to patient safety as it could delay diagnosis, allowing the cancer to progress. Therefore, our study selected thresholds to maximize sensitivity, at the cost of reduced specificity, by using the highest malignant lesion ADC as the cutoff (resulting in no false negatives; 100% sensitivity) for safer adoption into clinical workflows. However, it is important to acknowledge false negatives are virtually unavoidable when applying data derived thresholds on 'new' datasets, as illustrated by cross-validation testing in our study. Use of a more conservative inflated ADC cutoff to keep sensitivity high may be warranted in clinical practice [e.g., a 10% inflated cutoff was proposed in the ECOG-ACRIN 6602 trial (10)]. Regarding ROI approach, we found similar biopsy reduction rates could be achieved using 2D, 3D or hotspot lesion ADC values in MRI detected lesions. However, hotspot ADC measures were systematically lower and required a lower cutoff value vs. 2D to achieve equal performance. While the choice of ROI approach did not appear to affect the performance based on lesion size (big or small), our data suggested that measuring the hotspot (lowest ADC region) of the lesion may incrementally improve diagnostic performance over 2D and 3D ROI approaches for NMEs, warranting further investigation in a

TABLE 6 Performance of different ADC measures using prespecified cutoffs.

		A6702 ADC Threshold (1.53 x 10^{-3} mm^2/s)						EUSOBI ADC Threshold (1.3 x 10^{-3} mm^2/s)					
		Sensitivity			Benign Biopsy Reduction Rate			Sensitivity			Benign Biopsy Reduction Rate		
ADC Map	ROI Technique	No.	Est.	(95% CI)	No.	Est.	(95% CI)	No.	Est.	(95% CI)	No.	Est.	(95% CI)
4 b-values	2D	27/30	90.0%	(73.5, 97.9%)	21/107	19.6%	(12.1, 27.1%)	21/30	70.0%	(50.6, 85.3%)	55/107	51.4%	(41.7, 61.1%)
	3D	28/30	93.3%	(77.9, 99.2%)	21/107	19.6%	(12.1, 27.1%)	22/30	73.3%	(54.1, 87.7%)	54/107	50.5%	(40.7, 60.2%)
	Hotspot	30/30	100.0%	(88.4, 100.0%)	12/107	11.2%	(4.8, 17.6%)	25/30	83.3%	(65.3, 94.4%)	39/107	36.4%	(27.2, 45.7%)
2 b-values	2D	30/30	100.0%	(88.4, 100.0%)	21/107	19.6%	(12.1, 27.1%)	21/30	70.0%	(50.6, 85.3%)	55/107	51.4%	(41.9, 60.9%)
	3D	29/30	96.7%	(82.8, 99.9%)	20/107	18.7%	(11.3, 26.1%)	21/30	70.0%	(50.6, 85.3%)	55/107	51.4%	(41.7, 61.1%)
	Hotspot	30/30	100.0%	(88.4, 100.0%)	12/107	11.2%	(4.8, 17.6%)	25/30	83.3%	(65.3, 94.4%)	37/107	34.6%	(25.1, 44.1%)

ADC, apparent diffusion coefficient; CI, confidence interval; Est., Estimate; ROI, region of interest.

larger cohort. While prior studies have found limited diagnostic value of DWI in NME lesions (29, 30), a focused ROI approach may improve diagnostic accuracy by emphasizing the tumor regions of highest cellularity (16, 31).

In addition to deriving optimal ADC cutoffs from this dataset, we also applied two pre-defined cutoffs to validate their performance for the various ADC measurement approaches. We evaluated the cutoff identified by ECOG-ACRIN A6702 multicenter trial (1.53 x 10^{-3} mm^2/s) that prioritizes high sensitivity, which we confirmed maintained very high sensitivity (90-100%) in our dataset. This cutoff worked best for the 2D and 3D ROI approaches, with 19.6% reduction in benign biopsies (similar to the data-derived ADC cutoff), but had relatively lower diagnostic performance when using hotspot ROI measures (achieving only 11.2% reduction in benign biopsies). On the other hand, the lower EUSOBI recommended cutoff (1.3 x 10^{-3} mm^2/s), which is being utilized in some active multicenter trials [e.g., DWIST (22)], achieved the best performance (36.4% reduction in benign biopsies and 83.3% sensitivity) using the hotspot ROI approach in our dataset because the EUSOBI cutoff led to much lower sensitivity when using the other ROI techniques (70.0% and 73.3% for 2D and 3D, respectively). Regarding the choice of b-values, fitting of DWI signal intensities using a monoexponential decay model to calculate ADC is more robust with a greater number of b-values (4 vs 2), but at the cost of longer acquisition time. Our results found no significant difference in ADC measures computed with 2 vs 4 b-values for any of the ROI approaches, with very similar diagnostic performance and reduction in biopsies. These results are consistent with previous studies that have investigated optimal b-value combinations for breast DWI (15, 18, 32) and support the use of a two-b-value combination of 0 and 800 sec/mm^2 for optimal efficiency.

It is well recognized that *in vivo* ADC measures are affected by the b-values used for ADC calculation (15, 33), and different scanner platforms may have varying degrees of bias due to gradient nonlinearity effects (34), while variations in other factors such as spatial resolution and field strength could introduce other effects. A strength of our study therefore was the standardized data collection, which was performed at a single institution where the MRI scanner and protocol were kept consistent over the study period. Furthermore, focused inclusion criterion of lesions detected by screening MRI only [as opposed to palpable lesions, incidental findings in cancer staging MRI exams, or problem solving exams included in prior studies (15, 17, 35)] was used to generate a unique clinically-relevant dataset to evaluate impact on MRI screening performance.

This study has several limitations. While our study focused primarily on mean ADC values for each lesion, and we did not find any advantages of using other histogram metrics, more comprehensive radiomics based measures may further improve lesion characterization. Our study followed consensus guidelines for breast tumor ADC calculation (18) and did not explore alternate b value schemas (such as using maximum b > 800 s/mm^2 or non-zero minimum b), which would likely result in different ADC values and optimal cutoffs. Noise was not considered when calculating ADC, which may have introduced bias. Additionally, taking into account that the MRI signal follows a Rician distribution, especially in low signal-to-noise scenarios, could help make ADC estimates more consistent across different protocols, scanner hardware, and centers (36). Also, only monoexponential modeling was used for ADC map generation, while more advanced non-Gaussian, multi-compartment and other DWI modeling techniques may better characterize tissue microstructure and improve performance (37, 38). However, utilizing such advanced DWI models in breast screening applications is challenging due to limitations on scan time, small lesion sizes, and variable image quality of breast DWI in general (as ADC can be more robust to noise effects compared to other modeling parameters) (39, 40). Furthermore, all the measurements were performed offline using custom built software tools. Testing these ADC measurement techniques on clinical workstations may be needed to facilitate safe and real world implementation of DWI. Implementation of novel correction techniques during acquisition such as for gradient nonlinearity effects (34) and EPI distortions (41) are areas of future investigation to improve accuracy of ADC measures. Lastly, a larger sample size may be needed to identify subtle differences in diagnostic performance, particularly between the 2D and 3D techniques since most lesions were not measurable on multiple slices.

In conclusion, our findings demonstrate that unnecessary biopsies can be avoided for screening breast MRI exams while maintaining high sensitivity using a variety of ROI methods and b-value combinations for lesion ADC measurement. 2D, 3D and hotspot ROI approaches achieved similar rates of benign biopsy reduction using data derived ADC thresholds, which require further validation. The prespecified ECOG-ACRIN A6702 ADC cutoff worked best for 2D and 3D ROIs, whereas the lower EUSOBI cutoff was better suited for hotspot measures. Choice of ADC cutoff depends on ROI approach and preferred performance tradeoffs (biopsy reduction vs sensitivity). Shorter acquisitions with two b-values (0, 800 s/mm^2) might be sufficient, as the diagnostic performance was similar to that of the longer four b-value (0, 100, 600, 800 s/mm^2) acquisition. For safe and successful clinical integration of DWI to reduce biopsies, any of these ROI approaches and/or cutoffs could be applied but they need to be held consistent to achieve optimal diagnostic performance.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Fred Hutch Cancer Center Hutch IR#7339. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because of the retrospective nature of the study.

Author contributions

DB: Conceptualization, Data curation, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. DH: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – review & editing, Writing – original draft. AW: Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – review & editing. IL: Data curation, Investigation, Project administration, Resources, Writing – review & editing. HR: Conceptualization, Supervision, Writing – review & editing, Writing – original draft. SP: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The study was funded by NIH/NCI grants R01CA207290 and R01CA190299.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2024.1437506/full#supplementary-material.

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. (2024) 74:12–49. doi: 10.3322/caac.21820

2. Kuhl C, Weigel S, Schrading S, Arand B, Bieling H, König R, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *J Clin Oncol*. (2010) 28:1450–7. doi: 10.1200/JCO.2009.23.0839

3. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. *Jama*. (2012) 307:1394–404. doi: 10.1001/jama.2012.388

4. Riedl CC, Luft N, Bernhart C, Weber M, Bernathova M, Tea MK, et al. Triple-modality screening trial for familial breast cancer underlines the importance of magnetic resonance imaging and questions the role of mammography and ultrasound regardless of patient mutation status, age, and breast density. *J Clin Oncol*. (2015) 33:1128–35. doi: 10.1200/JCO.2014.56.8626

5. Sardanelli F, Podo F, Santoro F, Manoukian S, Bergonzi S, Trecate G, et al. Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. *Invest Radiol*. (2011) 46:94–105. doi: 10.1097/RLI.0b013e3181f3fcdf

6. Dong H, Kang L, Cheng S, Zhang R. Diagnostic performance of dynamic contrast-enhanced magnetic resonance imaging for breast cancer detection: An update meta-analysis. *Thorac Cancer*. (2021) 12:3201–7. doi: 10.1111/1759-7714.14187

7. Partridge SC, McDonald ES. Diffusion weighted magnetic resonance imaging of the breast: protocol optimization, interpretation, and clinical applications. *Magn Reson Imaging Clin N Am*. (2013) 21:601–24. doi: 10.1016/j.mric.2013.04.007

8. Vermoolen MA, Kwee TC, Nievelstein RA. Apparent diffusion coefficient measurements in the differentiation between benign and Malignant lesions: a systematic review. *Insights Imaging*. (2012) 3:395–409. doi: 10.1007/s13244-012-0175-y

9. Iima M, Partridge SC, Le Bihan D. Six DWI questions you always wanted to know but were afraid to ask: clinical relevance for breast diffusion MRI. *Eur Radiol*. (2020) 30:2561–70. doi: 10.1007/s00330-019-06648-0

10. Rahbar H, Zhang Z, Chenevert TL, Romanoff J, Kitsch AE, Hanna LG, et al. Utility of diffusion-weighted imaging to decrease unnecessary biopsies prompted by breast MRI: A trial of the ECOG-ACRIN cancer research group (A6702). *Clin Cancer Res*. (2019) 25:1756–65. doi: 10.1158/1078-0432.CCR-18-2967

11. Youn I, Biswas D, Hippe DS, Winter AM, Kazerouni AS, Javid SH, et al. Diagnostic performance of point-of-care apparent diffusion coefficient measures to reduce biopsy in breast lesions at MRI: clinical validation. *Radiology*. (2024) 310: e232313. doi: 10.1148/radiol.232313

12. Spick C, Pinker-Domenig K, Rudas M, Helbich TH, Baltzer PA. MRI-only lesions: application of diffusion-weighted imaging obviates unnecessary MR-guided breast biopsies. *Eur Radiol*. (2014) 24:1204–10. doi: 10.1007/s00330-014-3153-6

13. D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. *ACR BI-RADS® Atlas, breast imaging reporting and data system*. Reston, VA, USA: American College of Radiology (2013).

14. Nogueira L, Brandão S, Matos E, Nunes RG, Loureiro J, Ferreira HA, et al. Diffusion-weighted imaging: determination of the best pair of b-values to discriminate breast lesions. *Br J Radiol*. (2014) 87:20130807. doi: 10.1259/bjr.20130807

15. McDonald ES, Romanoff J, Rahbar H, Kitsch AE, Harvey SM, Whisenant JG, et al. Mean apparent diffusion coefficient is a sufficient conventional diffusion-weighted MRI metric to improve breast MRI diagnostic performance: results from the ECOG-ACRIN cancer research group A6702 diffusion imaging trial. *Radiology*. (2021) 298:60–70. doi: 10.1148/radiol.2020202465

16. Bickel H, Pinker K, Polanec S, Magometschnigg H, Wengert G, Spick C, et al. Diffusion-weighted imaging of breast lesions: Region-of-interest placement and different ADC parameters influence apparent diffusion coefficient values. *Eur Radiol*. (2017) 27:1883–92. doi: 10.1007/s00330-016-4564-3

17. Bickel H, Clauser P, Pinker K, Helbich T, Biondic I, Brkljacic B, et al. Introduction of a breast apparent diffusion coefficient category system (ADC-B) derived from a large multicenter MRI database. *Eur Radiol*. (2023) 33:5400–10. doi: 10.1007/s00330-023-09675-0

18. Baltzer P, Mann RM, Iima M, Sigmund EE, Clauser P, Gilbert FJ, et al. Diffusion-weighted imaging of the breast-a consensus and mission statement from the EUSOBI International Breast Diffusion-Weighted Imaging working group. *Eur Radiol*. (2020) 30:1436–50. doi: 10.1007/s00330-019-06510-3

19. Tanner EOSJE. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *J Chem Phys*. (1965) 42:288–92. doi: 10.1063/1.1695690

20. Biswas D, Hippe DS, Wang Y, DelPriore MR, Zečević M, Scheel JR, et al. Accelerated breast diffusion-weighted imaging using multiband sensitivity encoding with the CAIPIRINHA method: clinical experience at 3 T. *Radiol Imaging Cancer*. (2022) 4:e210063. doi: 10.1148/rycan.210063

21. Lo Gullo R, Sevilimedu V, Baltzer P, Le Bihan D, Camps-Herrero J, Clauser P, et al. A survey by the European Society of Breast Imaging on the implementation of breast diffusion-weighted imaging in clinical practice. *Eur Radiol*. (2022) 32:6588–97. doi: 10.1007/s00330-022-08833-0

22. Shin HJ, Lee SH, Park VY, Yoon JH, Kang BJ, Yun B, et al. Diffusion-weighted magnetic resonance imaging for breast cancer screening in high-risk women: design and imaging protocol of a prospective multicenter study in korea. *J Breast Cancer*. (2021) 24:218–28. doi: 10.4048/jbc.2021.24.e19

23. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. (1986) 42:121–30. doi: 10.2307/2531248

24. Huang FL. Using cluster bootstrapping to analyze nested data with a few clusters. *Educ Psychol Meas*. (2018) 78:297–318. doi: 10.1177/0013164416678980

25. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*. (1934) 26:404–13. doi: 10.1093/biomet/26.4.404

26. Baltzer A, Dietzel M, Kaiser CG, Baltzer PA. Combined reading of Contrast Enhanced and Diffusion Weighted Magnetic Resonance Imaging by using a simple sum score. *Eur Radiol*. (2016) 26:884–91. doi: 10.1007/s00330-015-3886-x

27. Lin CX, Tian Y, Li JM, Liao ST, Liu YT, Zhan RG, et al. Diagnostic value of multiple b-value diffusion-weighted imaging in discriminating the Malignant from benign breast lesions. *BMC Med Imaging*. (2023) 23:10. doi: 10.1186/s12880-022-00950-y

28. Song SE, Park EK, Cho KR, Seo BK, Woo OH, Jung SP, et al. Additional value of diffusion-weighted imaging to evaluate multifocal and multicentric breast cancer detected using pre-operative breast MRI. *Eur Radiol*. (2017) 27:4819–27. doi: 10.1007/s00330-017-4898-5

29. Marino MA, Avendano D, Sevilimedu V, Thakur S, Martinez D, Lo Gullo R, et al. Limited value of multiparametric MRI with dynamic contrast-enhanced and diffusion-weighted imaging in non-mass enhancing breast tumors. *Eur J Radiol*. (2022) 156:110523. doi: 10.1016/j.ejrad.2022.110523

30. Avendano D, Marino MA, Leithner D, Thakur S, Bernard-Davila B, Martinez DF, et al. Limited role of DWI with apparent diffusion coefficient mapping in breast lesions presenting as non-mass enhancement on dynamic contrast-enhanced MRI. *Breast Cancer Res*. (2019) 21:136. doi: 10.1186/s13058-019-1208-y

31. Arponen O, Sudah M, Masarwah A, Taina M, Rautiainen S, Könönen M, et al. Diffusion-weighted imaging in 3.0 tesla breast MRI: diagnostic performance and tumor characterization using small subregions vs. Whole tumor regions of interest. *PloS One*. (2015) 10(10):e0138702. doi: 10.1371/journal.pone.0141833

32. Partridge SC, Steingrimsson J, Newitt DC, Gibbs JE, Marques HS, Bolan PJ, et al. Impact of alternate b-value combinations and metrics on the predictive performance and repeatability of diffusion-weighted MRI in breast cancer treatment: results from the ECOG-ACRIN A6698 trial. *Tomography*. (2022) 8:701–17. doi: 10.3390/tomography8020058

33. Bogner W, Gruber S, Pinker K, Grabner G, Stadlbauer A, Weber M, et al. Diffusion-weighted MR for differentiation of breast lesions at 3.0 T: how does selection of diffusion protocols affect diagnosis? *Radiology*. (2009) 253(2):341–51. doi: 10.1148/radiol.2532081718

34. Malyarenko DI, Newitt DC, Amouzandeh G, Wilmes LJ, Tan ET, Marinelli L, et al. Retrospective correction of ADC for gradient nonlinearity errors in multicenter breast DWI trials: ACRIN6698 multiplatform feasibility study. *Tomography*. (2020) 6:86–92. doi: 10.18383/j.tom.2019.00025

35. Clauser P, Krug B, Bickel H, Dietzel M, Pinker K, Neuhaus VF, et al. Diffusion-weighted imaging allows for downgrading MR BI-RADS 4 lesions in contrast-enhanced MRI of the breast to avoid unnecessary biopsy. *Clin Cancer Res*. (2021) 27:1941–8. doi: 10.1158/1078-0432.CCR-20-3037

36. Walker-Samuel S, Orton M, McPhail LD, Robinson SP. Robust estimation of the apparent diffusion coefficient (ADC) in heterogeneous solid tumors. *Magn Reson Med*. (2009) 62:420–9. doi: 10.1002/mrm.22014

37. Cho GY, Moy L, Kim SG, Baete SH, Moccaldi M, Babb JS, et al. Evaluation of breast cancer using intravoxel incoherent motion (IVIM) histogram analysis: comparison with Malignant status, histological subtype, and molecular prognostic factors. *Eur Radiol*. (2016) 26:2547–58. doi: 10.1007/s00330-015-4087-3

38. Honda M, Le Bihan D, Kataoka M, Iima M. Diffusion kurtosis imaging as a biomarker of breast cancer. *BJR Open*. (2023) 5:20220038. doi: 10.1259/bjro.20220038

39. Mendez AM, Fang LK, Meriwether CH, Batasin SJ, Loubrie S, Rodríguez-Soto AE, et al. Diffusion breast MRI: current standard and emerging techniques. *Front Oncol*. (2022) 12:844790. doi: 10.3389/fonc.2022.844790

40. Iima M, Honda M, Sigmund EE, Ohno Kishimoto A, Kataoka M, Togashi K. Diffusion MRI of the breast: Current status and future directions. *J Magn Reson Imaging*. (2020) 52:70–90. doi: 10.1002/jmri.26908

41. Rodríguez-Soto AE, Fang LK, Holland D, Zou J, Park HH, Keenan KE, et al. Correction of artifacts induced by B(0) inhomogeneities in breast MRI using reduced-field-of-view echo-planar imaging and enhanced reversed polarity gradient method. *J Magn Reson Imaging*. (2021) 53(5):1581–91. doi: 10.1002/jmri.27566

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact