Diagnostic Performance and Accuracy of Strain Elastography for BI-RADS Category 4 Lesions among Asian Females

Meng Dong¹, Boyuan Xing¹, Bingyi Zhang¹, Xingdong Xu² and Quan Zhou²

¹Department of Ultrasound, The People’s Hospital of China Three Gorges University, The First People’s Hospital of Yichang, Hubei, China
²Department of General Surgery, The People’s Hospital of China Three Gorges University, The First People’s Hospital of Yichang, Hubei, China
³Department of Gynaecology and Obstetrics, The People’s Hospital of China Three Gorges University, The First People’s Hospital of Yichang, Hubei, China

ABSTRACT
This work aimed to assess the diagnostic performance of strain elastography (elasticity score) and its accuracy for breast BI-RADS category 4 lesions. Online databases including Cochrane Library, Embase, PubMed, and Web of Science were searched for eligible articles published prior to March 10, 2022. The pooled effect indicators including sensitivity, specificity, positive and negative likelihood ratios (PLR, NLR), the area under the summary receiver operating characteristic (SROC) curve, and diagnostic odds ratio (DOR) were utilised to assess the strain elastography’s performance in diagnosing BI-RADS category 4 lesions. Subgroup analyses and meta-regressions were used to explore the potential sources of heterogeneity, and the Deeks’ funnel plot asymmetry test was used to detect any publication bias. The literature search yielded 11 studies involving 5028 BI-RADS category 4 lesions (including 1809 malignant lesions). The recruited lesions were all from Asian females. The pooled sensitivity, specificity, PLR, NLR, and DOR with the 95% confidence intervals were 0.68 (0.66-0.71), 0.83 (0.82-0.85), 3.36 (2.45–4.60), 0.32 (0.21-0.49), and 12.11 (7.46-19.65), respectively. The area under the SROC curve was 0.85. No significant publication bias was detected. Taken together, strain elastography had suboptimal sensitivity but desirable specificity for the accurate diagnosis of BI-RADS category 4 lesions among Asian females, which can help avoid unnecessary biopsies and reduce patient anxiety.

Key Words: BI-RADS category 4 lesions, Strain elastography, Meta-analysis.

INTRODUCTION
Breast cancer ranked as the most commonly diagnosed malignancy worldwide in 2020, which could further bring somatopsychic disorders among its female sufferers. Early detection and diagnosis, and appropriate, timely treatment are paramount and significantly helpful in improving their quality of life and survival rate. According to the BI-RADS guideline for breast lesions, BI-RADS category 4 lesions are subdivided into low (4A, likelihood: >2% to ≤10%), moderate (4B, likelihood: >10% to ≤50%), and high (4C, likelihood: >50% to <95%) suspicious ones in 2013.

Since the wide range of likelihood of malignancy for such category lesions and morphological overlap between malignant and benign lesions, the conventional B-mode US presents a relatively low specificity in the diagnosis of breast malignancies. US elastography, developed in 1991 by Ophir et al. and rapidly applied as a new clinically diagnostic modality, has been confirmed in the American College of Radiology’s consensus on the 5th BI-RADS guideline for its better lesion characterisation and risk stratification of the breast. This technology is sensitive to tissue stiffness and can provide a non-invasive evaluation of the stiffness of breast lesions based on two-dimensional ultrasound, namely, providing effective, complementary information support for precise diagnosis.

Previous meta-analyses have assessed the diagnostic accuracy of strain elastography or shear wave for all BI-RADS categories of breast lesions or non-palpable ones. These studies conclude that strain elastography could serve as an effective technique to complement the B-mode US in predicting malignant breast lesions and thus, reduce biopsies for benign breast lesions. This work is currently the first meta-analysis for the assessment of the most controversial BI-RADS category 4 lesions.

lesions with strain elastography. This study aimed to assess the
diagnostic performance of strain elastography (elasticity score)
and its accuracy for BI-RADS category 4 lesions.

METHODOLOGY

This meta-analysis was conducted as per the PRISMA reporting
guidelines updated in 2020. A systematic literature search in
online databases (Embase, Cochrane Library, PubMed, Web of
Science) was performed by two authors independently. There
was no limitation on the start date for targeting publications,
and the end date was March 10, 2022. The search strategy
included the Medical Subject Heading terms: elasticity imaging
techniques, elasticity, elastogram, tissue elasticity imaging,
sonoelastography, elastography, strain elastography, real time
elastography breast lesions, and BI-RADS 4. The two authors
independently judged whether a potentially eligible article
could be used as per the pre-established inclusion and exclu-
sion criteria. Any discrepancies in data collection were resolved
via discussion with a third author.

Publications that had assessment test(s) for the diagnostic accu-
uracy of strain elastography (elasticity score) for BI-RADS cate-
gory 4 lesions, and pathological investigation of the specimen(s) via biopsy or surgery (serving as the reference standard
for diagnostic confirmation), and provided direct or indirect
data about false positive (FP), true positive (TP), true negative
(TN), false negative (FN) rates or their number of cases, were
included. Literature was excluded if presented as a review, a
correspondence or conference article, or duplicate ones, or not
written in English, or had less than 50 samples (breast lesions).

The extracted data from each eligible publication included the
first author’s surname, publication year, study design type,
country, the patients’ mean age, the number of lesions, mean
lesion size, excitation method for strain elastography, the refer-
ce standard for confirming a diagnosis, the cut-off value for
diagnosing a malignancy, and numbers of TP, FP, TN, and FN
breast lesions. The methodological quality of a publication was
evaluated by the same authors independently using a revised
QUADAS tool for assessing the quality of the diagnostic accu-

RESULTS

As shown in Figure 1, the initial literature search in the four
databases yielded 1138 articles/studies. After removing dupli-
cate studies, the remaining 875 articles received serious
screening and checked for relevance concerning their titles/abs-
abstracts; 731 articles were excluded owing to their failure to meet
the inclusion criteria. The other 144 full-text ones were
assessed for eligibility. Eventually, 11 articles were selected for
the meta-analysis.

As shown in Table I, 5028 BI-RADS category 4 lesions (including
1809 malignant lesions) were recruited in the meta-analysis. All
the recruited lesions were from Asian females. The studies were
conducted across China (seven articles), Korea (three articles),
and Japan (one article). There were seven prospective and four
retrospective studies.
Five brands of US machines equipped with strain elastography function were used in the 11 included studies, namely, Hitachi from Japan (five studies), General Electric from the USA (three studies), Siemens from Germany (one study), Mindray from China (one study), and Samsung from Korea (one study). Siemens’s excitation method for strain elastography in the study by Han et al. was cardiovascular/respiratory pulsation, while manual compression was used for excitation in the other 10 studies. Six studies had a sample size (number of lesions) of less than 200, and the other five had more than 200 sample lesions.

As shown in Figure 2, the risk of bias in the recruited studies came from patient selection, flow and timing. A high risk of bias was detected in four studies regarding patient selection and two studies regarding flow and timing. An unclear risk of bias was present in one study regarding patient selection and two studies regarding flow and timing. An unclear risk of bias was present in one study regarding patient selection and two studies regarding flow and timing. A high risk of bias was detected in four studies regarding patient selection and two studies regarding flow and timing. An unclear risk of bias was present in one study regarding patient selection and two studies regarding flow and timing. Of the 11 included articles, four were retrospective studies and three did not mention the use of consecutive patient enrollment. Besides, exclusion of palpable BI-RADS category 4 lesions and assessment of non-palpable breast lesions alone in the studies by Cho et al. and Yi et al. could introduce a risk of bias.
In this meta-analysis, the positive Spearman correlation coefficient was 0.718 and the p-value was 0.013, which suggested the existence of a threshold effect and consequently, the potential presence of heterogeneity. As illustrated in Figures 3A-D and Figure 4, the pooled sensitivity, specificity, PLR, NLR, and DOR with the 95% confidence intervals were 0.68 (0.66–0.71) (p<0.001, I²=97.1%), 0.83 (0.82–0.85) (p<0.001, I²=97.9%), 3.36 (2.45–4.60) (p<0.001, I²=90.3%), 0.32 (0.21–0.49) (p<0.001, I²=96.9%), and 12.11 (7.46–19.65) (p<0.001, I²=81.0%), respectively. As shown in Figure 5, the area under the SROC curve was 0.85, indicating a relatively high diagnostic accuracy of strain elastography in differentiating benign BI-RADS category 4 lesions from malignant ones. The inconsistency (I² value) for sensitivity and specificity was 97.1% and 97.9% respectively (both p<0.001) reflecting a significant heterogeneity within the included studies.

In terms of the substantial heterogeneity, subgroup analyses were performed based on country (China or others), study design type (prospective or retrospective), the number of lesions (>200 or <200), and excitation method (cardiovascular/respiratory pulsation or manual compression), respectively (shown in Table II). Besides, as shown in Table III, the meta-regressions confirmed that none of the four subgroup factors was the source of heterogeneity. With the symmetrical shape of the funnel plot, the absence of publication bias was suggested by the p-value of 0.19 (Figure 6).
DISCUSSION

BI-RADS category 4 lesions are usually suspicious enough to warrant a recommendation for biopsies, especially in the absence of typical imaging for malignancy. A prospective multicentre study in China reported that less than half of BI-RADS category 4 lesions (925/2049) were pathologically confirmed to be malignant. In a clinical practice, over-diagnosis of breast lesions brings about overtreatment and collateral increase in patients’ medical expenses, as well as unnecessary anxiety for the patients. Hence, more accurate non-invasive diagnostic methods are required to raise the specificity of diagnosing BI-RADS category 4 lesions and thus, reduce the requirement for biopsies. Encouragingly, shear wave and strain elastography can provide a better quantitative as well as qualitative assessment of breast masses. Their addition to conventional B-mode US could increase the specificity from 0.27 to 0.76 without loss in sensitivity (p <0.001). Several studies have recently demonstrated that the strain elastography-derived elasticity scores above 3 are highly associated with malignant breast lesions (p <0.05). Hao et al. proposed that category 4A lesions, when assessed with an elasticity score of 1, could be downgraded to category 3 (without the intervention of biopsies); a short-term follow-up or tissue-based diagnosis was recommended for the category 4A lesions with a score of 2 or 3. In this meta-analysis, 11 eligible studies were finally included, and the recruited lesions were all from Asian females. The pooled sensitivity and specificity were 0.68 and 0.83, respectively, and the area under the SROC was 0.85, which manifested a relatively high diagnostic value for the category 4 lesions. Another similar meta-analysis focused on the elasticity score for all breast lesions and achieved consistent but superior sensitivity (0.77) and specificity (0.88). Most category 2 and 3 lesions are benign and the majority of category 5 lesions are considered malignant. Thus, the reliable diagnostic outcome based on elasticity scores from these lesions could explain the slight gap in sensitivity and specificity. FP and FN results are unavoidable when applying strain elastography, partly due to the overlap between the characteristics of benign and malignant category 4 lesions. Increased stiffness of lesions owing to calcification in fibroadenoma, atypical ductal hyperplasia, and sclerosing adenosis, would yield a higher FP rate. The papillary proliferation of papillomas in the mammary duct and secondary changes (haemorrhage, fibrosis, or infarction) could affect the elasticity scores and FP results. FN results mainly come from mucinous carcinoma and ductal carcinoma in situ. The mucin-rich structure of the former could cause lower elasticity scores. The unobvious invasive and desmoplastic reactions of the latter at its early stage could cause lower elasticity scores. The pooled PLR and NLR in this meta-analysis were 3.36 and 0.32. This means that breast cancer patients had 3.36 times greater odds (95% CI 2.45-4.60) of receiving an elasticity score of above 3 than healthy controls, and the error rate was 32% when confirming the TN cases through breast elasticity scoring. DOR serves as a single indicator for the accuracy of a diagnostic test, and its higher value indicates a better
The discriminatory test performance. The DOR in the present meta-analysis was 12.11 (95% CI 7.46-19.65), indicating a relatively high diagnostic value of elasticity score in category 4 lesions.

Exploration of heterogeneity was an essential issue when running a meta-analysis. The Spearman correlation coefficient was 0.718 (p <0.05) in this meta-analysis, which indicated the presence of a threshold effect (a primary cause of heterogeneity). Afterwards, through subgroup analysis, it was found that the retrospective studies showed relatively higher sensitivity and DOR but lower specificity than the prospective studies. The studies conducted in China possessed a better diagnostic performance than those performed in other countries. Besides, the groups with a smaller sample size presented higher accuracy in diagnosing BI-RADS category 4 lesions through strain elastography compared with the groups with a larger sample size. Notably, the only study, whose strain elastography was excited by cardiovascular/respiratory pulsation, had higher sensitivity but lower specificity in contrast to the pooled corresponding values of the other ten studies (through manual compression). Further large-size studies concentrating on the latter three subgroup factors are needed to confirm these findings. Meta-regressions were then performed from four aspects (country, study design type, number of lesions, and excitation method), and none of them significantly affected the heterogeneity.

Without a doubt, several limitations exist in this meta-analysis. Firstly, in the very beginning it aimed to focus on category 4 lesions among females, but the recruited lesions in the included studies were all from oriental Asian females (Chinese, Korean and Japanese) and not from South Asia. In addition, over half of lesions (3300/5028) were from Chinese females, and there were obvious differences between Asian and non-Asian women in breast masses. Therefore, the findings might not be extrapolated to other Asian or non-Asian female populations. More multicentric and prospective studies involving patients from various countries or regions are required for further confirmation. Secondly, substantial heterogeneity was detected in this meta-analysis, but subgroup analyses and meta-regressions did not detect any sources of heterogeneity. Thirdly, the data about mean breast lesion size were provided in only two included studies, and the influence of lesion size on elasticity degree was not studied in this meta-analysis. Hence, the influence of the size or mean size of category 4 lesion on elasticity score in the diagnostic accuracy test demands further investigation.

**CONCLUSION**

This meta-analysis indicated that the diagnostic accuracy of strain elastography has unsatisfactory sensitivity but desirable specificity for BI-RADS category 4 lesions among Asian females, which can reduce unnecessary biopsies and decrease patient anxiety and medical cost.

**FUNDING:**
This work was supported by the Yichang Science and Technology Bureau Project, China (Grant No. A21-2-029).

**ACKNOWLEDGEMENT:**
The authors would like to sincerely acknowledge the assistance of the combined faculty of the Department of Ultrasound, The People’s Hospital of China Three Gorges University and The First People’s Hospital of Yichang, for providing guidance and necessary critique during the study period. In particular, we are grateful to Prof. Chaojun Yang and Jun Yang for their intensive role in this work. We are also grateful to the library management for providing us with the electronic databases that made this work possible.

**COMPETING INTEREST:**
The authors declared no competing interest.

**AUTHORS’ CONTRIBUTION:**
MD, BX: Designed the study, searched the literature, collected and analysed the data.
XX, QZ, BZ: Collected the data and provided interpretation of the results.
BZ: Drafted the manuscript.
All authors revised and approved the final version of the manuscript.

**REFERENCES**

Diagnostic accuracy of strain elastography for breast lesions


