REVIEW

Ultrasonographic differentiation of malignant from benign breast lesions: a meta-analytic comparison of elasticity and BIRADS scoring

Gelareh Sadigh · Ruth C. Carlos · Colleen H. Neal ·
Ben A. Dwamena

Abstract There has been controversy regarding the accuracy of breast ultrasound elastography (USE) compared to conventional B-mode Ultrasound (USB). The purpose of this study was to conduct a direct comparative effectiveness analysis of USB versus USE or their combination in differentiating breast lesions through systematically reviewing recent literature. An extensive literature search of PubMed and other medical and general purpose databases from inception through August 2011 was conducted. Published studies that reported a direct comparison of the diagnostic performance of USE, using elasticity score versus USB, using breast imaging reporting and data system (BIRADS) for characterization of focal breast lesions were included. Summary diagnostic performance measures were assessed for each of the tests and their combination using bivariate generalized linear mixed modeling. The two tests were combined as: (1) conjunctive, where the outcome of the combination of tests is positive only if both test results are positive; (2) disjunctive, where the outcome of a combination of tests is negative only if both tests are negative. Twenty nine studies provided relevant information on 5,511 breast masses (2,065 cancers, 3,446 benign lesions). Sensitivity of USB, USE, and their conjunctive and disjunctive combinations were 96% (95% credible interval (CrI), 93–98%), 79% (95% CrI, 74–83%), 73% (95% CrI, 67–78%), and 99% (95% CrI, 98–99%), respectively. Specificity of USB, USE, and their conjunctive and disjunctive combinations were 70% (95% CrI, 55–83%), 88% (95% CrI, 82–92%), 97% (95% CrI, 95–99%), and 56% (95% CrI, 43–69%), respectively. The application of USE as a single test is not superior to USB alone. However, in low risk patients it is recommended to perform an USE following a positive USB result to decrease the rate of unnecessary biopsies.

Keywords Breast lesion · Breast ultrasound · Elastography · Malignancy

G. Sadigh (✉)
Department of Radiology, University of Michigan Medical Center, 1500 E. Medical Center Dr, B1 132 H Taubman Center, Ann Arbor, MI 48109-5302, USA
e-mail: sadighg@umich.edu

R. C. Carlos
Division of Magnetic Resonance Imaging, Department of Radiology, University of Michigan Medical Center, Ann Arbor, MI, USA

R. C. Carlos · C. H. Neal · B. A. Dwamena
VA Ann Arbor Healthcare System, Ann Arbor, MI, USA

C. H. Neal
Division of Breast Imaging, Department of Radiology, University of Michigan Medical Center, Ann Arbor, MI, USA

B. A. Dwamena
Division of Nuclear Medicine, Department of Radiology, University of Michigan Medical Center, Ann Arbor, MI, USA

Abbreviations

AUC	Area under curve
USB	B-mode ultrasound
USE	Breast ultrasound elastography
BIRADS	Breast imaging reporting and data system
CrI	Credible interval
ES	Elasticity score
FN	False negatives
FP	False positives
I^2	Inconsistency index

LR	Likelihood ratios
QUADAS	Quality assessment of diagnostic accuracy studies
SROC	Summary receiver operating characteristic
TN	True negatives
TP	True positives

Introduction

Breast ultrasound elastography (USE) is a new technique of ultrasonic imaging that has shown effectiveness for detection of malignancy within breast lesions. USE provides information about the mechanical properties of tissue such as elasticity and strain and maps it into color images [1–4]. Elasticity is the tendency of a tissue to resume the original size and shape; while strain is the level of change in size or shape in response to external compression (stress) [4]. Each pixel of the image is assigned one of 256 specific colors and demonstrates the magnitude of tissue strain depending on physiological and pathological changes in breast structure [3, 5]. Harder tissues such as malignancy may result in decreased strain and are shown in blue, while softer tissues will reflect increased strain and are shown in red [3]. Normal breast tissue which reflects average strain is shown in Green [3].

The color image is superimposed on B-mode ultrasound (USB) image for a better recognition of the relationship between the strain distribution and the anatomical borders of the lesion [3, 4, 6]. This information is further interpreted by evaluating the color pattern in a hypoechoic lesion (e.g., within lesion borders on USB image), and in the surrounding breast tissue [3]. A 1 to 5 scale elasticity score (ES) is assigned to each image based on its overall pattern, with the harder tissues (e.g. breast cancer) showing higher elasticity scores [3].

The diagnostic accuracy of elasticity scoring has already been investigated in several previous studies [2, 3, 5, 7–36] and a previous meta-analysis [37]. Our prior meta-analysis has shown a sensitivity of 79% and a specificity of 89% for use of ES in differentiating benign and malignant lesions [38]. Further, individual studies have reported that USE alone may increase the specificity of breast ultrasound in the characterization of breast lesions and potentially decrease unnecessary biopsies of benign breast lesions [1, 7, 9, 16, 39]. Only seven studies have reported the diagnostic performance of combination of USE (using ES) and USB, using breast imaging reporting and data system (BIRADS) [2, 8, 17, 22, 27, 33, 40]. Five of them have shown an improvement in the specificity of USB [8, 22, 27, 33, 40]; while four have reported an increase in sensitivity

[2, 17, 27, 33]. However, evidence is lacking for a meta-analysis that directly evaluates the diagnostic performance of USB compared to USE alone or its combination with USB. We performed a meta-analysis of studies that reported a direct comparison of ES with BIRADS in differentiating breast lesions, according to Cochrane Diagnostic Test Accuracy Review Working Group guidelines [41]. We further, evaluated the diagnostic performance of combination of USE and USB compared to USB, using specific statistical methods [42] on the same database, representing a significant expansion of our previous work [38].

Methods

Criteria for considering studies for this review

Types of studies

All analytical studies reporting a direct comparison of elasticity score alone with BIRADS in differentiation of focal breast lesions that were published in full text were considered for eligibility. No language restriction was used.

Participants

Study participants were patients who had breast symptoms or an abnormal clinical breast examination, breast US or mammography. There was no age restriction for the study participants.

Index test

Breast USE was the index test. Only papers in which a 5-point scale elasticity score according to Itoh et al. [3] was calculated were included. Lesions with ES of 4 and 5 were considered malignant, while the other ES were grouped as benign lesions.

Comparator test

Conventional USB was the comparator. USB images were reported according to BIRADS categories [43]. Lesions with BIRADS categories of 4 and 5 were considered malignant, while the other categories were grouped as benign lesions.

Target condition

The index test is used to differentiate benign from malignant breast lesions.

Reference standards

Histopathological (core biopsy or surgical biopsy) or cytological (fine needle aspiration) confirmation of breast lesion is the reference standard.

Search methods for identification of studies

For the purpose of this study, we used the search result of our prior meta-analysis [38].

Electronic searches

Electronic searches of PubMed, EMBASE, ISI Web of Knowledge and Cochrane database from inception through to August 22, 2011 were performed without any constraints. We used relevant text words and Medical Subject Heading terms that included *breast* combined with *sono-elastography*, *elastosonography*, *elastography*, *elasticity imaging* and *strain imaging*.

Searching other resources

Reference lists from identified studies were manually scanned to identify other relevant studies.

Data collection and analysis

Selection of studies

Two authors independently conducted the literature search. A list of articles meeting the inclusion criteria based on abstracts was complied, and these articles were retrieved in full text. Two reviewers independently reviewed the list of full texts for inclusion. Discrepancies were discussed and resolved upon agreement on a final set of studies.

Data extraction and management

Data, extracted by two reviewers (GS, BAD), included patient characteristics (number, gender, mean age), lesion palpability, technical characteristics of USE, the reference standard, and the study results for USE and USB (number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN)).

Assessment of methodological quality

Two reviewers (GS, BAD) independently assessed the methodological quality of included studies, using 11 items of the extensively validated Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist [44]. Disagreements were resolved by consensus.

Statistical analysis and data synthesis

TP, FP, TN and FN for USE and USB were extracted directly from the source literature, where possible. Otherwise, values were calculated from the data provided. If the study did not report results for USB, we requested original data by directly corresponding with authors or principal investigators. Summary sensitivity and specificity and 95% credible interval (CrI)—which is the Bayesian analog to confidence interval—for USE and USB were calculated using bivariate generalized linear mixed modeling (a random effects model) [45, 46]. We used Bayesian Markov chain Monte Carlo simulation with non-informative hyperpriors and implemented with the WinBUGS program interfaced with STATA. Summary positive and negative likelihood ratios (LR) were calculated from the model estimates. LRs can be interpreted as follows: a LR of 0 excludes disease, a LR of infinity (∞) excludes normality and a LR of 1 means no change in likelihood of disease. For the diagnostic information to have high probability of altering clinical management, a likelihood ratio greater than 10 or less than 0.1 would be required for a positive or negative test result, respectively. Moderate informational value can be achieved with likelihood ratios of 5–10 and 0.1–0.2; likelihood ratios of 2.0–5.0 and 0.2–0.5 indicate very little informational value [47].
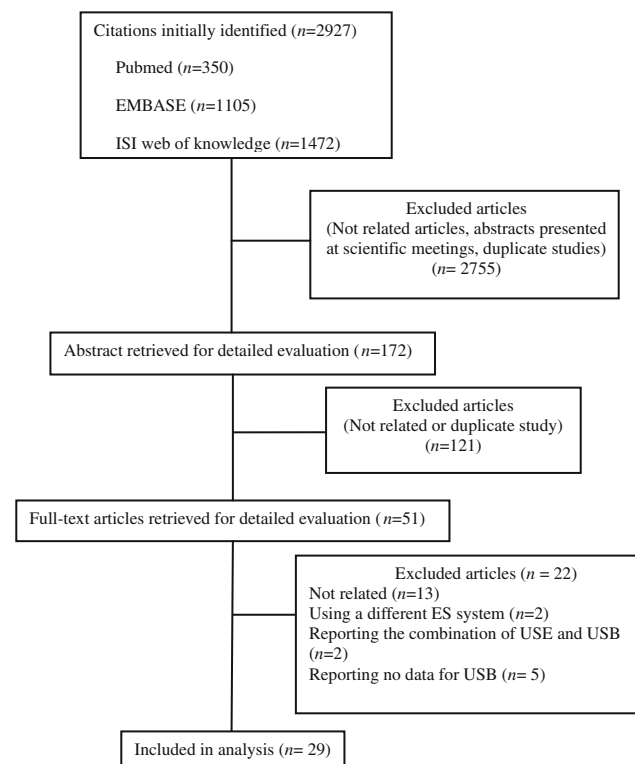

Fig. 1 Literature search and selection schema. *USB* B-mode ultrasound, *USE* ultrasound elastography

The data were graphically displayed in summary receiver operating characteristic (SROC) curve with summary operating points for sensitivity and specificity on the curves embellished with 95% confidence region. We used the Rutter and Gastonis version of formulas for constructing SROC curve [48].

Further, we assessed the diagnostic accuracy (sensitivity, specificity, and LRs and area under SROC curve) of combination of USE and USB, using the method described by Ament et al. [42]. We combined the two tests based on two different positivity criterion: (1) conjunctive, where the outcome of the combination of tests is positive only if both test results are positive; in all other cases the outcome of the combination of tests would be negative; (2) disjunctive, where the outcome of a combination of tests is negative only if both tests are negative; in all other cases, the outcome of the combination of tests would be positive.

Publication bias was assessed using Deeks' funnel plot asymmetry test [49]. Heterogeneity between studies was assessed by using the I^2 statistics. I^2 values range between 0 and 100%, where 0% indicates no observed heterogeneity and values greater than 50% may be considered to indicate substantial heterogeneity [50]. All statistical analyses were performed with the user-written "midas" module for STATA, version 11 (Stata Corp., College Station, Texas) [51].

Results

The literature search from our previous study yielded 2,927 articles, of which 172 were reviewed in abstract, and from them 51 were further reviewed in full text (Fig. 1). Of these, 29 studies were eligible for inclusion. Of the excluded studies, thirteen did not fully meet the inclusion

Table 1 Baseline characteristics of included studies

Included articles	No of patients[a]	No of lesions[a]	Palpable (%)	Women (%)	Mean age (years)	Reference standard
Navarro et al. [7]	116	124	43	100	46	Histo/cyto
Bartolotta et al. [34]	72	84	44	100	47	Histo/cyto
Satake et al. [30]	104	115	68	100	55	Histo
Fischer et al. [9]	201	201	0	100	53	Histo
Lee et al. [8]	278	315	ND	100	47	Histo
Yerli et al. [10]	71	78	ND	100	52	Histo
Parajuly et al. [17]	150	170	ND	100	45	Histo/cyto
Ciurea et al. [18]	92	101	ND	100	48	Histo/cyto
Wojcinski et al. [16]	779	779	ND	100	54	Histo/cyto
Hatzung et al. [2]	97	97	55	98	55	Histo/cyto
Chung et al. [19]	120	120	0	100	47	Histo
Raza et al. [20]	175	188	ND	100	47	Histo
Thomas et al. [12]	227	227	ND	100	54	Histo
Regini et al. [5]	110	101	ND	100	51	Histo/cyto
Moon et al. [57]	140	140	0	100	46	Histo
Shaefer et al. [23]	193	193	ND	100	54	Histo/cyto
Sohn et al. [22]	267	281	ND	100	ND	Histo
Scaperrotta [28]	278	293	0	100	53	Histo/cyto
Tan et al. [26]	415	535	71	100	39	Histo/cyto
Vanhoutte [25]	59	65	8	100	51	Histo/cyto
Cho et al. [29]	100	100	40	100	46	Histo
Zhu et al. [24]	112	139	57	100	46	Histo
Tardivan et al. [31]	114	122	ND	100		Histo/cyto
Thomas et al. [15]	50	50	62	100	49	Histo
Zhi et al. [27]	232	296	ND	100	42	Histo
Thomas et al. [14]	300	300	57	100	51	Histo/cyto
Thomas et al. [13]	108	108	0	100	54	Histo/cyto
Itoh et al. [3]	111	111	ND	100	49	Histo/cyto
Giuseppetti et al. [32]	82	91	ND	100	55	Histo/cyto

Cyto cytopathology, *Histo* histolopathogy, *ND* not described, *No* number

[a] Number of patients and lesions are the final number that was included in the analysis in each study

criteria, two had used 4-point scale elasticity score [52, 53], measured different from the classification introduced by Itoh et al. [3], two did not report data on USE performance as a single test [33, 40], and five did not report data on USB performance [11, 21, 35, 36, 54].

Characteristics of included studies

Table 1 summarizes the clinical characteristics of patients and their breast lesions, and the reference standard used in all the included meta-analyses. Included studies were published between July 2005 and May 2011 in peer-reviewed journals. The mean age of included patients ranged from 39 to 55 years. Overall, our analysis included 29 studies with information on 5,153 patients and 5,511 breast masses. Of included lesions, 2,065 were malignant and 3,446 were benign. Except for Parajuly et al. [17], which used acoustic vibration source for strain measurement, the rest of the studies used freehand compression

elastography probes. All the studies that reported the correlation method for strain measurement, had used the combined autocorrelation method (CAM) [3].

Methodological quality of included studies

Appendix Fig. 1 in Supplementary material summarizes the frequency of each study quality indicator across the studies. All studies used appropriate reference standard(s) for verification, and 97% of them explained withdrawals. On the other hand, none of the studies clearly stated whether the reference standard interpretation was performed without the knowledge of index test results.

Diagnostic performance of USE alone

The summary test operating measures for USE alone were: sensitivity of 79% (95% confidence interval (CrI), 74–83%), specificity of 88% (95% CrI, 82–92%) (Fig. 2;
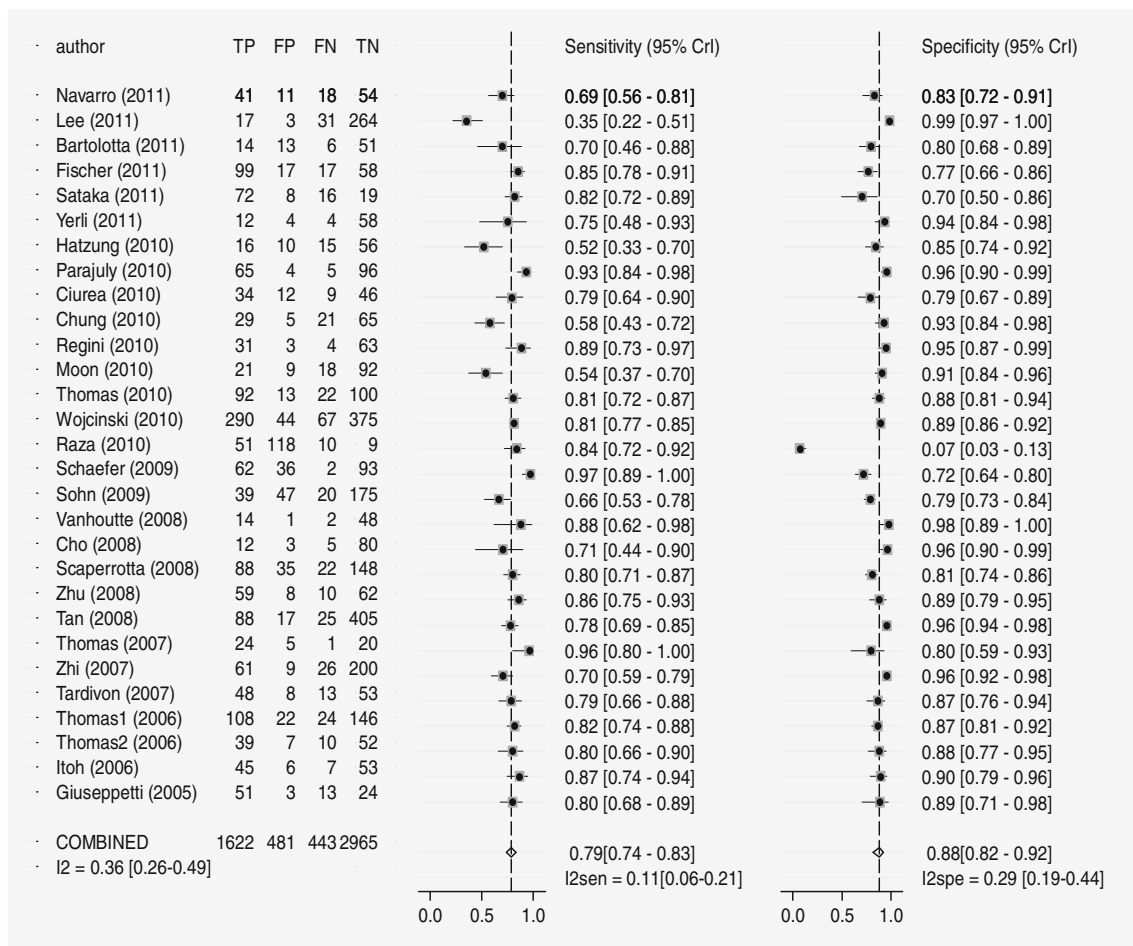
author	TP	FP	FN	TN	Sensitivity (95% CrI)	Specificity (95% CrI)
Navarro (2011)	41	11	18	54	0.69 [0.56 - 0.81]	0.83 [0.72 - 0.91]
Lee (2011)	17	3	31	264	0.35 [0.22 - 0.51]	0.99 [0.97 - 1.00]
Bartolotta (2011)	14	13	6	51	0.70 [0.46 - 0.88]	0.80 [0.68 - 0.89]
Fischer (2011)	99	17	17	58	0.85 [0.78 - 0.91]	0.77 [0.66 - 0.86]
Sataka (2011)	72	8	16	19	0.82 [0.72 - 0.89]	0.70 [0.50 - 0.86]
Yerli (2011)	12	4	4	58	0.75 [0.48 - 0.93]	0.94 [0.84 - 0.98]
Hatzung (2010)	16	10	15	56	0.52 [0.33 - 0.70]	0.85 [0.74 - 0.92]
Parajuly (2010)	65	4	5	96	0.93 [0.84 - 0.98]	0.96 [0.90 - 0.99]
Ciurea (2010)	34	12	9	46	0.79 [0.64 - 0.90]	0.79 [0.67 - 0.89]
Chung (2010)	29	5	21	65	0.58 [0.43 - 0.72]	0.93 [0.84 - 0.98]
Regini (2010)	31	3	4	63	0.89 [0.73 - 0.97]	0.95 [0.87 - 0.99]
Moon (2010)	21	9	18	92	0.54 [0.37 - 0.70]	0.91 [0.84 - 0.96]
Thomas (2010)	92	13	22	100	0.81 [0.72 - 0.87]	0.88 [0.81 - 0.94]
Wojcinski (2010)	290	44	67	375	0.81 [0.77 - 0.85]	0.89 [0.86 - 0.92]
Raza (2010)	51	118	10	9	0.84 [0.72 - 0.92]	0.07 [0.03 - 0.13]
Schaefer (2009)	62	36	2	93	0.97 [0.89 - 1.00]	0.72 [0.64 - 0.80]
Sohn (2009)	39	47	20	175	0.66 [0.53 - 0.78]	0.79 [0.73 - 0.84]
Vanhoutte (2008)	14	1	2	48	0.88 [0.62 - 0.98]	0.98 [0.89 - 1.00]
Cho (2008)	12	3	5	80	0.71 [0.44 - 0.90]	0.96 [0.90 - 0.99]
Scaperrotta (2008)	88	35	22	148	0.80 [0.71 - 0.87]	0.81 [0.74 - 0.86]
Zhu (2008)	59	8	10	62	0.86 [0.75 - 0.93]	0.89 [0.79 - 0.95]
Tan (2008)	88	17	25	405	0.78 [0.69 - 0.85]	0.96 [0.94 - 0.98]
Thomas (2007)	24	5	1	20	0.96 [0.80 - 1.00]	0.80 [0.59 - 0.93]
Zhi (2007)	61	9	26	200	0.70 [0.59 - 0.79]	0.96 [0.92 - 0.98]
Tardivon (2007)	48	8	13	53	0.79 [0.66 - 0.88]	0.87 [0.76 - 0.94]
Thomas1 (2006)	108	22	24	146	0.82 [0.74 - 0.88]	0.87 [0.81 - 0.92]
Thomas2 (2006)	39	7	10	52	0.80 [0.66 - 0.90]	0.88 [0.77 - 0.95]
Itoh (2006)	45	6	7	53	0.87 [0.74 - 0.94]	0.90 [0.79 - 0.96]
Giuseppetti (2005)	51	3	13	24	0.80 [0.68 - 0.89]	0.89 [0.71 - 0.98]
COMBINED	1622	481	443	2965	0.79 [0.74 - 0.83]	0.88 [0.82 - 0.92]
I2 = 0.36 [0.26-0.49]					I2sen = 0.11 [0.06-0.21]	I2spe = 0.29 [0.19-0.44]

Fig. 2 Forest plot of studies reporting elasticity score shows individual estimated sensitivities and specificities of the studies evaluated in the meta-analysis, as well as pooled values (*open diamond*), with corresponding 95% credible intervals (*in brackets*).

The *broken black line* represents the pooled estimates of sensitivity and specificity. *FN* false-negative, *FP* false-positive, *TN* true-negative, *TP* true-positive, *95% CrI* 95% credible interval

Table 2). The positive and negative likelihood ratios (LR) for USE were 6.71 (95% CrI, 4.60–10.20) and 0.24 (95% CrI, 0.19–0.30), respectively. This translates into a moderately informative test, where exclusion and confirmation of breast cancer using the test alone is not possible [38].

Figure 3 shows the resulting SROC curve with summary operating points for sensitivity and specificity on the curves. The summary area under the curve was 91% (95% confidence region, 89–93%), compatible with a good test accuracy [55]. The inconsistency index (I^2) for heterogeneity was 36% (95% CrI, 26–49%). When assessing I^2 for sensitivity

and specificity analysis separately, the index was 11% (95% CrI, 6–21%) for sensitivity and 29% (95% CrI, 19–44%) for specificity. Funnel plot and linear regression showed no evidence of publication bias.

Diagnostic performance of USB alone

USB summary estimate of sensitivity and specificity were 96% (95% CrI, 93–98%), and 70% (95% CrI, 55–83%), respectively (Fig. 4; Table 2). The pooled positive and negative LRs for USB alone were 3.10 (95% CI,

Table 2 Diagnostic and clinical performance of elasticity score, B-mode ultrasound and their combination

	Elasticity score	B-mode ultrasound	Conjunctive combination	Disjunctive combination
Sensitivity, % (95% credible interval)	79 (74–83)	96 (93–98)	73 (67–79)	99 (98–99)
Specificity, % (95% credible interval)	88 (83–92)	70 (55–83)	97 (95–98)	59 (44–68)
Positive likelihood ratio, (95% credible interval)	6.71 (4.60–10.20)	3.10 (2.12–5.14)	26.20 (16.00–48.68)	2.41 (1.79–3.05)
Negative likelihood ratio, (95% credible interval)	0.24 (0.19–0.30)	0.06 (0.04–0.10)	0.28 (0.22–0.34)	0.02 (0.01–0.02)
Area under SROC curve, % (95% confidence region)	91 (89–93)	92 (90–94)	98 (96–98)	91 (89–93)

Fig. 3 Summary receiver operator characteristics curves including a summary operating point for sensitivity and specificity (*green diamond*) and a 95% confidence region (*gray square*) for elasticity score (*left upper*), BIRADS (*right upper*), conjunctive combination of the two tests (*left lower*), and disjunctive combination of the two tests (*right lower*). The individual *circles* around each study number (observed data) describe the sample size weight of the individual studies. *AUC* area under curve; *SEN* sensitivity, *SPE* specificity
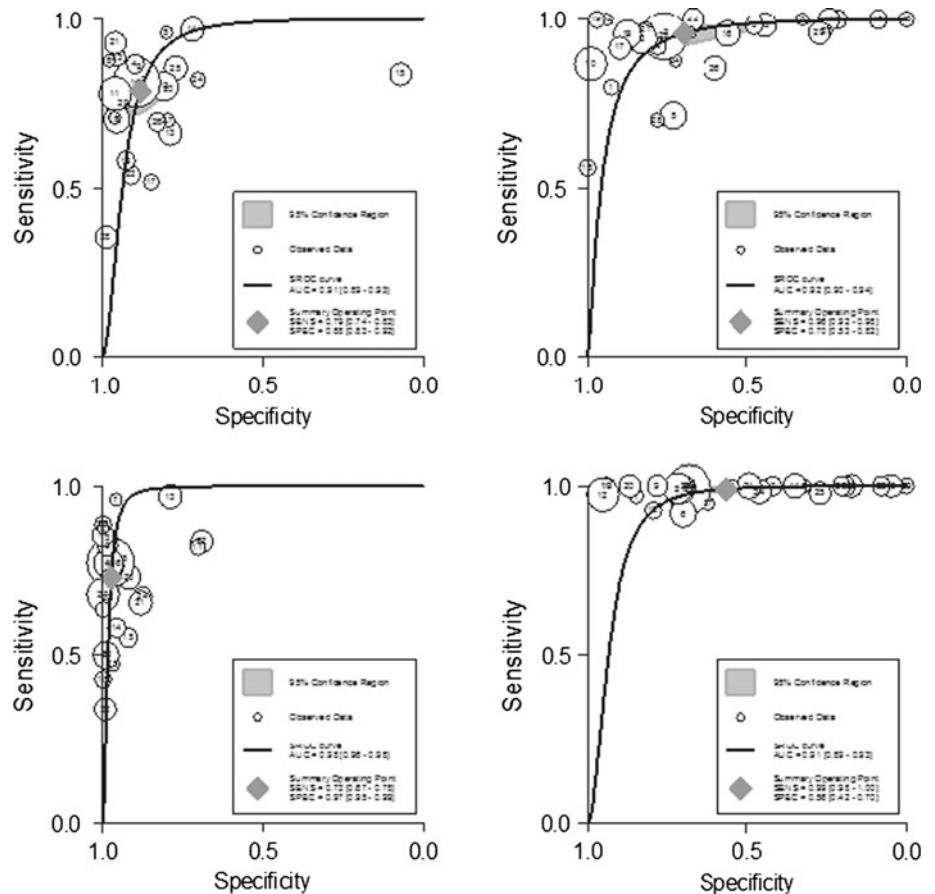
2.12–5.14) and 0.06 (95% CI, 0.04–0.10), respectively. This finding would be interpreted as a negative USB is capable of excluding malignancy within a breast lesion; while a positive USB is slightly informative and requires additional testing for confirmation of malignancy [47]. The summary area under the curve for USB was 92% (95% confidence region, 90–94%), compatible with a good test accuracy (Fig. 3) [55]. The I^2 for heterogeneity was 62% (95% CrI, 49–75%). When assessing I^2 for sensitivity and specificity analysis separately, the index was 35% (95% CrI, 20–55%) for sensitivity and 51% (95% CrI, 37–67%) for specificity.

Diagnostic performance of combination of USB and USE

Conjunctive positivity criterion

We further evaluated the diagnostic performance of combination of USB and USE using conjunctive positivity criterion. The analysis demonstrated that the summary sensitivity for combination of the tests was 73% (95% CrI,

67–79%), and the summary specificity was 97% (95% CrI, 95–98%) (Fig. 5; Table 2). The summary positive and negative LRs for the combination of the tests were 26.20 (95% CrI, 16.00–48.68) and 0.28 (95% CrI, 0.22–0.34), respectively, showing that the positive test is strongly capable of confirming the disease; while the negative test is as informative as USE alone, and requires further test for ruling out the cancer. The summary area under the curve for USB was 98% (95% confidence region, 96–98%), translated into a better detection performance compared to USE or USB alone (Fig. 3) [55, 56]. The I^2 for heterogeneity was 42% (95% CrI, 30–58%). When assessing I^2 for sensitivity and specificity analysis separately, the index was 13% (95% CrI, 7–24%) for sensitivity and 36% (95% CrI, 23–54%) for specificity.

Disjunctive positivity criterion

Assessing the combination of the two tests using the disjunctive positivity criterion showed that the summary sensitivity, specificity, positive LR and negative LR for combination of tests were 99 (95% CrI, 98–99%), 59%
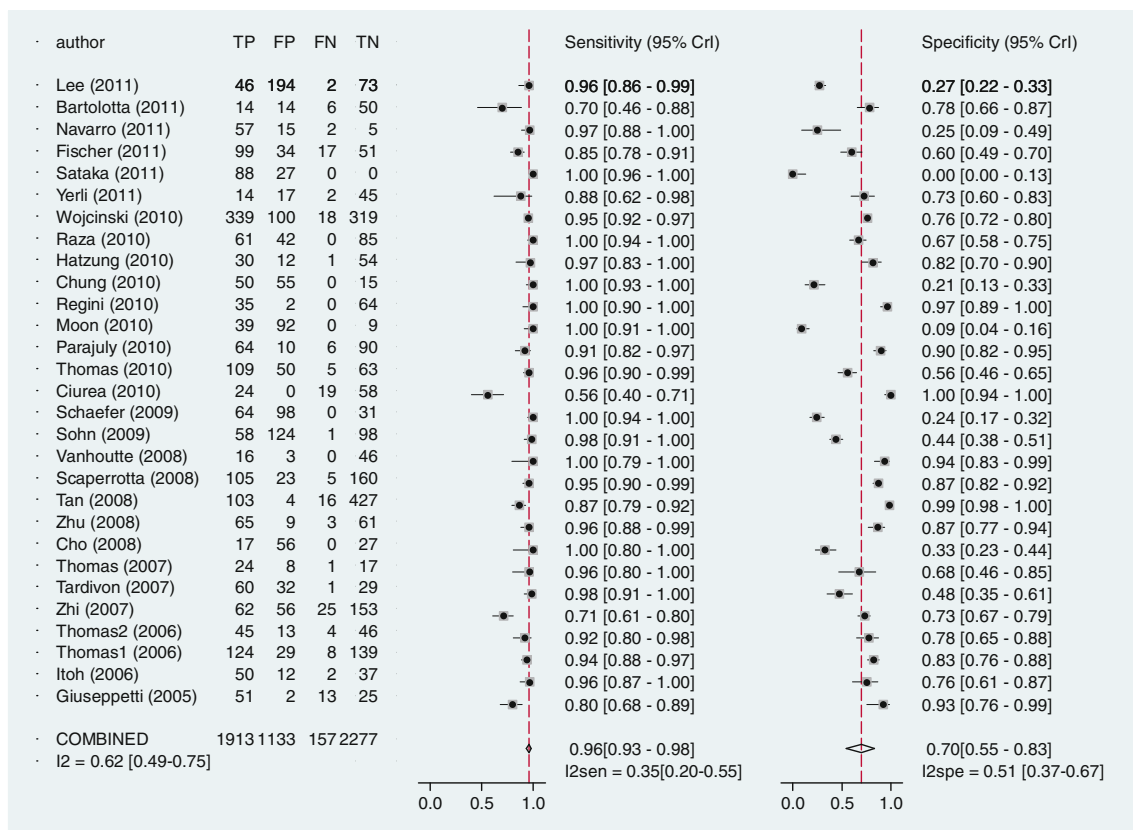

author	TP	FP	FN	TN	Sensitivity (95% CrI)	Specificity (95% CrI)
Lee (2011)	46	194	2	73	0.96 [0.86 - 0.99]	0.27 [0.22 - 0.33]
Bartolotta (2011)	14	14	6	50	0.70 [0.46 - 0.88]	0.78 [0.66 - 0.87]
Navarro (2011)	57	15	2	5	0.97 [0.88 - 1.00]	0.25 [0.09 - 0.49]
Fischer (2011)	99	34	17	51	0.85 [0.78 - 0.91]	0.60 [0.49 - 0.70]
Sataka (2011)	88	27	0	0	1.00 [0.96 - 1.00]	0.00 [0.00 - 0.13]
Yerli (2011)	14	17	2	45	0.88 [0.62 - 0.98]	0.73 [0.60 - 0.83]
Wojcinski (2010)	339	100	18	319	0.95 [0.92 - 0.97]	0.76 [0.72 - 0.80]
Raza (2010)	61	42	0	85	1.00 [0.94 - 1.00]	0.67 [0.58 - 0.75]
Hatzung (2010)	30	12	1	54	0.97 [0.83 - 1.00]	0.82 [0.70 - 0.90]
Chung (2010)	50	55	0	15	1.00 [0.93 - 1.00]	0.21 [0.13 - 0.33]
Regini (2010)	35	2	0	64	1.00 [0.90 - 1.00]	0.97 [0.89 - 1.00]
Moon (2010)	39	92	0	9	1.00 [0.91 - 1.00]	0.09 [0.04 - 0.16]
Parajuly (2010)	64	10	6	90	0.91 [0.82 - 0.97]	0.90 [0.82 - 0.95]
Thomas (2010)	109	50	5	63	0.96 [0.90 - 0.99]	0.56 [0.46 - 0.65]
Ciurea (2010)	24	0	19	58	0.56 [0.40 - 0.71]	1.00 [0.94 - 1.00]
Schaefer (2009)	64	98	0	31	1.00 [0.94 - 1.00]	0.24 [0.17 - 0.32]
Sohn (2009)	58	124	1	98	0.98 [0.91 - 1.00]	0.44 [0.38 - 0.51]
Vanhoutte (2008)	16	3	0	46	1.00 [0.79 - 1.00]	0.94 [0.83 - 0.99]
Scaperrotta (2008)	105	23	5	160	0.95 [0.90 - 0.99]	0.87 [0.82 - 0.92]
Tan (2008)	103	4	16	427	0.87 [0.79 - 0.92]	0.99 [0.98 - 1.00]
Zhu (2008)	65	9	3	61	0.96 [0.88 - 0.99]	0.87 [0.77 - 0.94]
Cho (2008)	17	56	0	27	1.00 [0.80 - 1.00]	0.33 [0.23 - 0.44]
Thomas (2007)	24	8	1	17	0.96 [0.80 - 1.00]	0.68 [0.46 - 0.85]
Tardivon (2007)	60	32	1	29	0.98 [0.91 - 1.00]	0.48 [0.35 - 0.61]
Zhi (2007)	62	56	25	153	0.71 [0.61 - 0.80]	0.73 [0.67 - 0.79]
Thomas2 (2006)	45	13	4	46	0.92 [0.80 - 0.98]	0.78 [0.65 - 0.88]
Thomas1 (2006)	124	29	8	139	0.94 [0.88 - 0.97]	0.83 [0.76 - 0.88]
Itoh (2006)	50	12	2	37	0.96 [0.87 - 1.00]	0.76 [0.61 - 0.87]
Giuseppetti (2005)	51	2	13	25	0.80 [0.68 - 0.89]	0.93 [0.76 - 0.99]
COMBINED	1913	1133	157	2277	0.96 [0.93 - 0.98]	0.70 [0.55 - 0.83]
I2 = 0.62 [0.49-0.75]					I2sen = 0.35 [0.20-0.55]	I2spe = 0.51 [0.37-0.67]

Fig. 4 Forest plot of studies shows individual estimated sensitivities and specificities for B-mode ultrasound in the studies evaluated in the meta-analysis, as well as pooled values (*open diamond*), with corresponding 95% credible intervals (*in brackets*). The broken black line represents the pooled estimates of sensitivity and specificity. *FN* false-negative, *FP* false-positive, *TN* true-negative, *TP* true-positive, *95% CrI* 95% credible interval
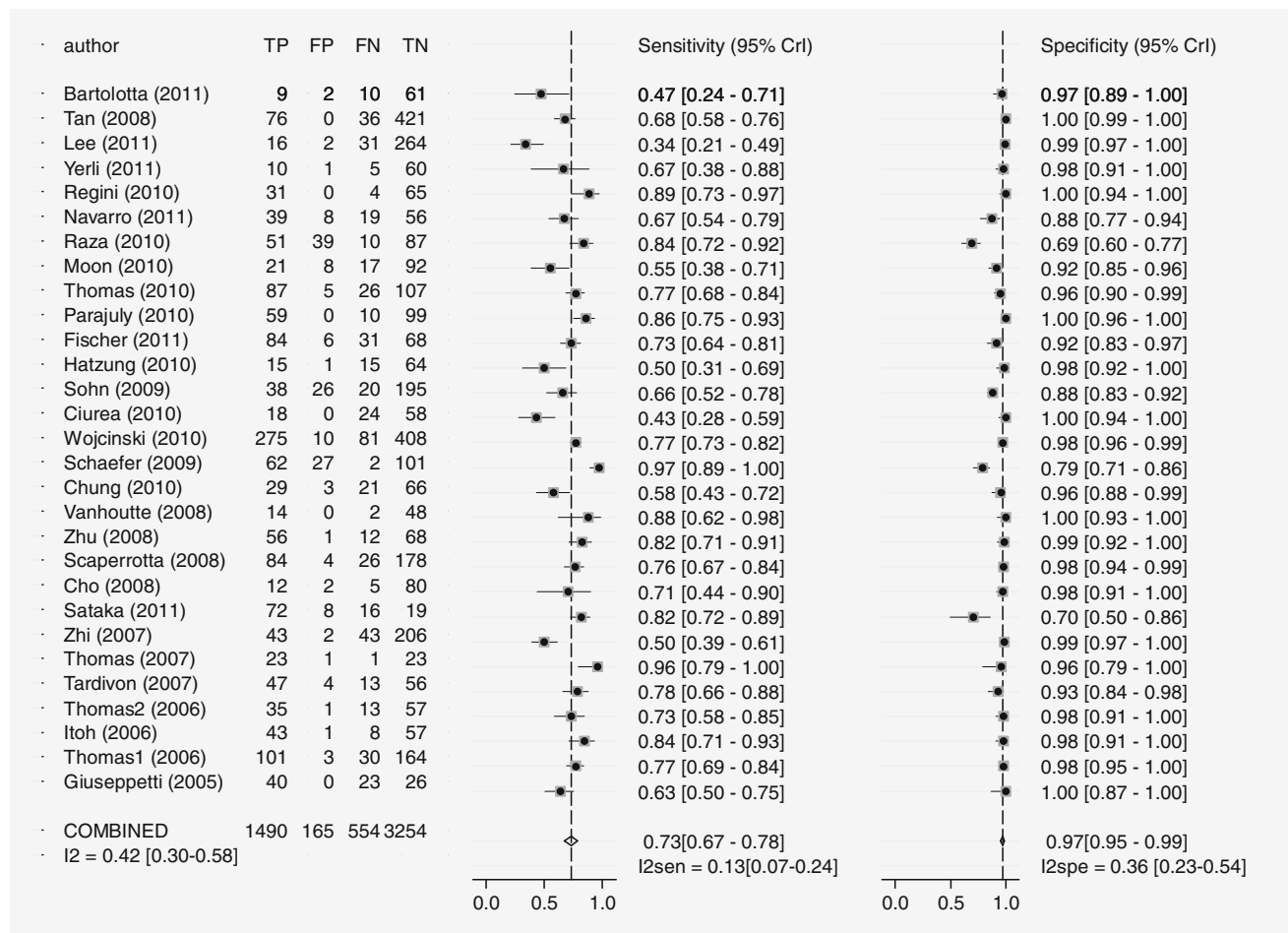
Fig. 5 Forest plot of studies shows individual estimated sensitivities and specificities for conjunctive combination of elastography and B-mode ultrasound of the studies evaluated in the meta-analysis, as well as pooled values (*open diamond*), with corresponding 95% credible intervals (*in brackets*). The *broken black line* represents the pooled estimates of sensitivity and specificity. *FN* false-negative, *FP* false-positive, *TN* true-negative, *TP* true-positive, *95% CrI* 95% credible interval

(95% CrI, 44–68%), 2.41 (95% CrI, 1.79–3.05), and 0.02 (95% CI, 0.01–0.02) (Fig 6; Table 2). This can be interpreted as a test that is nearly similar to USB alone; a negative test can exclude malignancy, while a positive test is small informative and requires further tests for confirmation. The summary area under the curve was 91% (95% confidence region, 89–93%) (Fig. 3). The I^2 for heterogeneity was 45% (95% CrI, 30–94%). When assessing I^2 for sensitivity and specificity analysis separately, the index was 5% (95% CrI, 0–43%) for sensitivity and 43% (95% CrI, 26–90%) for specificity.

Comparative effectiveness of USE versus USB

Figure 7 and Appendix Fig. 2 in supplementary material show the post-test probabilities of USE alone, USB alone and the combination of the two tests across a range of

disease prevalence (or pre-test probability) of breast cancer. According to Bayesian statistics, in patients with low pre-test probability of a disease, a high specificity is important [42]. Therefore, the conjunctive combination of the two tests which has the highest specificity (97%), compared to other options would be the best strategy to avoid unnecessary biopsies. On the other hand, for patients with high pre-test probability of disease, a high sensitivity is important, and therefore, the disjunctive combination of the two tests (sensitivity of 99%) or USB alone (sensitivity of 96%), would be the best test options.

Discussion

This study is the first comparative effectiveness meta-analysis of USE and USB and their combination on 5,511
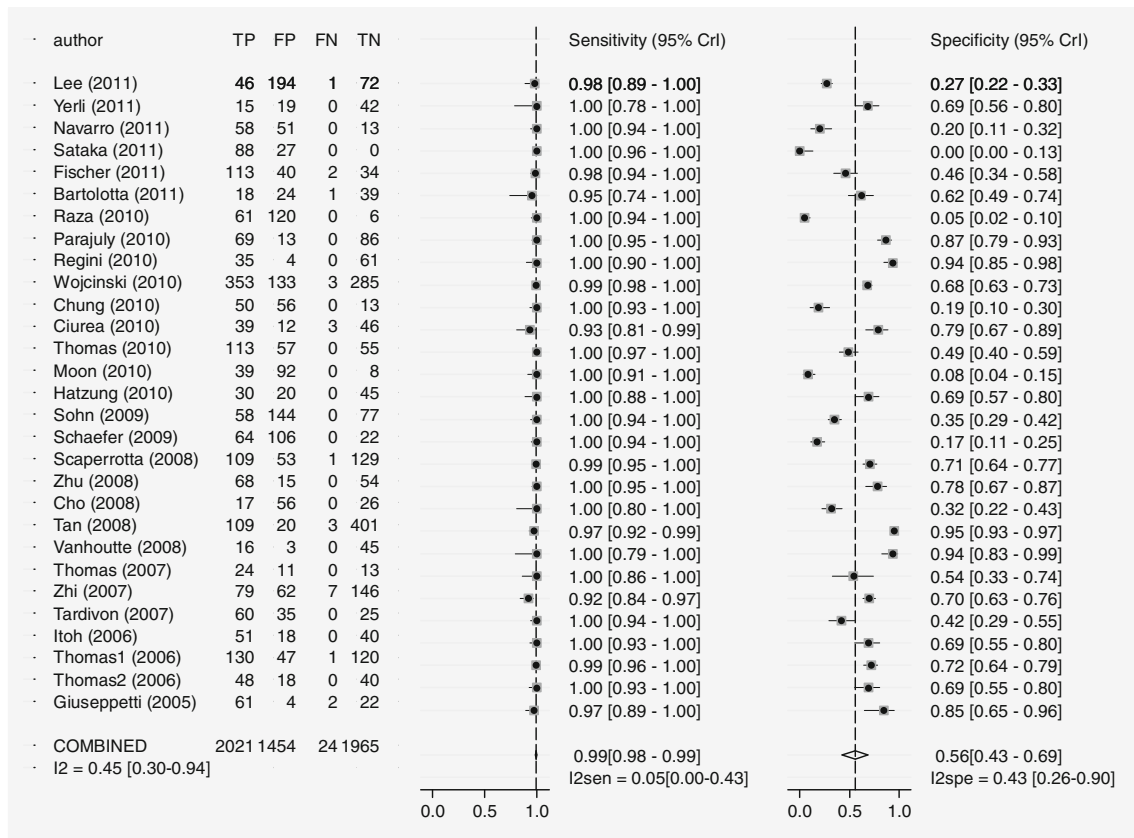
author	TP	FP	FN	TN	Sensitivity (95% CrI)	Specificity (95% CrI)
Lee (2011)	46	194	1	72	0.98 [0.89 - 1.00]	0.27 [0.22 - 0.33]
Yerli (2011)	15	19	0	42	1.00 [0.78 - 1.00]	0.69 [0.56 - 0.80]
Navarro (2011)	58	51	0	13	1.00 [0.94 - 1.00]	0.20 [0.11 - 0.32]
Sataka (2011)	88	27	0	0	1.00 [0.96 - 1.00]	0.00 [0.00 - 0.13]
Fischer (2011)	113	40	2	34	0.98 [0.94 - 1.00]	0.46 [0.34 - 0.58]
Bartolotta (2011)	18	24	1	39	0.95 [0.74 - 1.00]	0.62 [0.49 - 0.74]
Raza (2010)	61	120	0	6	1.00 [0.94 - 1.00]	0.05 [0.02 - 0.10]
Parajuly (2010)	69	13	0	86	1.00 [0.95 - 1.00]	0.87 [0.79 - 0.93]
Regini (2010)	35	4	0	61	1.00 [0.90 - 1.00]	0.94 [0.85 - 0.98]
Wojcinski (2010)	353	133	3	285	0.99 [0.98 - 1.00]	0.68 [0.63 - 0.73]
Chung (2010)	50	56	0	13	1.00 [0.93 - 1.00]	0.19 [0.10 - 0.30]
Ciurea (2010)	39	12	3	46	0.93 [0.81 - 0.99]	0.79 [0.67 - 0.89]
Thomas (2010)	113	57	0	55	1.00 [0.97 - 1.00]	0.49 [0.40 - 0.59]
Moon (2010)	39	92	0	8	1.00 [0.91 - 1.00]	0.08 [0.04 - 0.15]
Hatzung (2010)	30	20	0	45	1.00 [0.88 - 1.00]	0.69 [0.57 - 0.80]
Sohn (2009)	58	144	0	77	1.00 [0.94 - 1.00]	0.35 [0.29 - 0.42]
Schaefer (2009)	64	106	0	22	1.00 [0.94 - 1.00]	0.17 [0.11 - 0.25]
Scaperrotta (2008)	109	53	1	129	0.99 [0.95 - 1.00]	0.71 [0.64 - 0.77]
Zhu (2008)	68	15	0	54	1.00 [0.95 - 1.00]	0.78 [0.67 - 0.87]
Cho (2008)	17	56	0	26	1.00 [0.80 - 1.00]	0.32 [0.22 - 0.43]
Tan (2008)	109	20	3	401	0.97 [0.92 - 0.99]	0.95 [0.93 - 0.97]
Vanhoutte (2008)	16	3	0	45	1.00 [0.79 - 1.00]	0.94 [0.83 - 0.99]
Thomas (2007)	24	11	0	13	1.00 [0.86 - 1.00]	0.54 [0.33 - 0.74]
Zhi (2007)	79	62	7	146	0.92 [0.84 - 0.97]	0.70 [0.63 - 0.76]
Tardivon (2007)	60	35	0	25	1.00 [0.94 - 1.00]	0.42 [0.29 - 0.55]
Itoh (2006)	51	18	0	40	1.00 [0.93 - 1.00]	0.69 [0.55 - 0.80]
Thomas1 (2006)	130	47	1	120	0.99 [0.96 - 1.00]	0.72 [0.64 - 0.79]
Thomas2 (2006)	48	18	0	40	1.00 [0.93 - 1.00]	0.69 [0.55 - 0.80]
Giuseppetti (2005)	61	4	2	22	0.97 [0.89 - 1.00]	0.85 [0.65 - 0.96]
COMBINED	2021	1454	24	1965	0.99 [0.98 - 0.99]	0.56 [0.43 - 0.69]
I2 = 0.45 [0.30-0.94]					I2sen = 0.05 [0.00-0.43]	I2spe = 0.43 [0.26-0.90]

Fig. 6 Forest plot of studies shows individual estimated sensitivities and specificities for disjunctive combination of elastography and B-mode ultrasound of the studies evaluated in the meta-analysis, as well as pooled values (*open diamond*), with corresponding 95% credible intervals (*in brackets*). The *broken black line* represents the pooled estimates of sensitivity and specificity. *FN* false-negative, *FP* false-positive, *TN* true-negative, *TP* true-positive, *95% CrI* 95% credible interval

breast masses. Our study results demonstrated that ES can improve USB specificity (70 vs. 88%) at the cost of a drop in test sensitivity (96 vs. 78%). We further demonstrated that in patients with low risk of disease the conjunctive combination of the two tests would be the most useful test option, while in patients with high risk of breast cancer single USB test or the disjunctive combination of the two tests would be the best option.

Conventional USB, palpation and mammography are the three steps routinely performed in clinic for diagnosing a suspicious breast lesion. However, none of them alone or in combination with each other is able to differentiate malignancy and there is always the need to obtain biopsy or fine needle aspiration to confirm the diagnosis. A high percentage of these biopsies are benign [57], but may lead to increased patient's anxiety, and impose a burden of cost to health care system [58]. On the other hand, the risk of a missed malignancy in a non-palpable lesion always remains despite using USB and mammography [59]. Therefore, any improvement in medical technology which

can improve the diagnostic performance of these modalities is encouraged.

Breast USE has integrated the diagnostic ability of palpation into an ultrasound instrument with a compressive probe, and reflects the tissue stiffness (hardness) and elasticity in response to pressure; even in lesions that are not-palpable by hand. The 1–5 point scale ES introduced by Itoh et al. [3] has provided for standardized interpretation of elasticity images which may then be translated into a 1–5 point scale similar to BIRADS categories.

In this study, our results show that USE when used alone improves the specificity of USB. Compared to USB, a 19% increase in test specificity with ES may result in 17% decrease in sensitivity, with no overall improvement in test accuracy (area under curve (AUC), 91 vs. 92%) (Fig. 3). Considering the fact that USE is not currently reimbursed and overall test performance of USE alone compared to USB is similar, USE alone does not appear superior to USB to recommend its clinical use independent of USB.
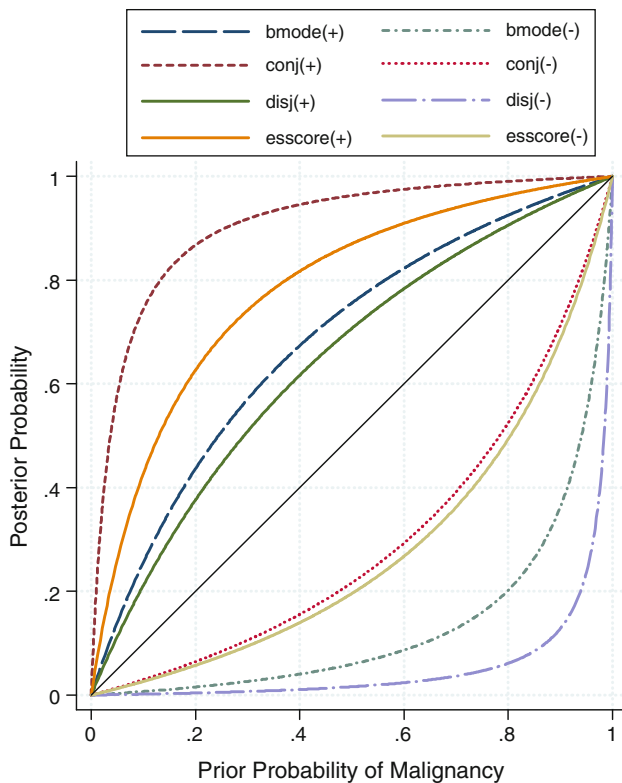
Fig. 7 Conditional probability curves after a positive and negative test result for US elastography (*esscore*), B-mode ultrasound (*bmode*), and their combination using conjunctive positivity criterion (*conj*), and disjunctive positivity criterion (*disj*). The *horizontal axis* shows the pre-test probability of malignancy within a breast lesion, and the *vertical axis* shows the post-test probability of malignancy

In evaluating the combination of USB and USE, the current study tested two combinations, the conjunctive combination and the disjunctive combination, as defined above. Our results demonstrate that the conjunctive combination of USE and USB, is capable of improving the specificity by 28% (compared to USB alone) at the cost of decreasing the sensitivity by 23%. However, the overall improvement in test accuracy (AUC, 97 vs. 92%), as well as the significantly high positive LR of 26.28 makes the test an ideal option to confirm the disease, resulting in a significant decrease in the number of unnecessary biopsies especially in low risk patients. In order to avoid unnecessary costs and to achieve the highest test efficiency, we recommend that all low risk patients first undergo USB, and only if the test was positive, a USE is performed as a supplemental modality to assist in decision to biopsy a lesion or follow it up.

The disjunctive combination of the two tests may result in a slight increase in test sensitivity compared to USB (99 vs. 96%), at the cost of a greater decrease in test specificity

(59 vs. 69%). Previous single-site studies (which we used as source publications) that reported the diagnostic performance of combination of USB and USE [2, 8, 17, 22, 27, 33, 40] used different definitions for what constituted a "combined USB/USE examination" (Table 3). Therefore, the literature shows heterogeneous improvement for a "combined USB/USE examination" compared to either test individually. The two studies [2, 17] that reported the disjunctive combination of the two tests fitting our definition have also shown improved sensitivity similar to our meta-analytic results. None of the prior studies conducted evaluation of conjunctive combination.

Limitations

Our study has limitations. First, the best cutoff score for determining benign or malignant lesions varied in different studies [3, 8, 20, 23, 57]. However, for the purpose of this study we extracted data with a cutoff score of 3 for all individual studies. Second, USE performs better in differentiation of malignancy within small lesions that are surrounded by a large amount of normal tissue, than for large lesions [18]. Scaperrotta et al. and Zhi et al. [21, 28] reported higher sensitivity but lower specificity of USE for lesions less than 1 cm in size. Giuseppetti et al. and Regini et al. [5, 32] reported an improvement in both sensitivity and specificity for lesions less than 2 cm. Since we did not have access to patient level data in the current study, it was not possible to evaluate the performance of USE or its combination with USB based on lesions size.

Implications

In summary, the application of USE as a single test is not superior to USB alone. USE improves specificity of conventional USB. However, this decrease in the number of unnecessary biopsies may be at the cost of increase in the number of missed cancers, with no change in overall diagnostic accuracy. However, in low risk patients, we recommend that USE be performed following a positive B-mode result. If both the USB and USE are positive, the patient should be referred for biopsy. Other patients with positive USB and negative USE could be evaluated with imaging follow-up which may serve to decrease the rate of benign biopsies. For high risk patients, we recommend that USB alone, rather than using USE alone or their combination, be used to evaluate breast masses. In these patients, if the USB is positive, we recommend further evaluation with biopsy.

Table 3 List of the studies that reported the diagnostic performance of USE and USB

	Rule for combining USE and USB	Sensitivity of combination compared to USB	Specificity of combination compared to USB
Hatzung et al. [2]	If USE OR USB is malignant ⇒ combination is malignant	↑	↓
	If USE AND USB are benign ⇒ combination is benign		
Parajuly et al. [17]	If USE OR USB is malignant ⇒ combination is malignant	↑	↔
	If USE AND USB are benign ⇒ combination is benign		
Lee et al. [8]	If ES = 1 ⇒ downgrade of BIRADS category for combination	↓	↑
	If ES = 2 or 3 ⇒ No change in BIRADS for combination		
	If ES = 4 or 5 ⇒ Upgrade of BIRADS category for combination		
Kadour et al. [33]	If unambiguous benign in USE AND USB ⇒ Combination score = 2	↑	↑
	If unambiguous malignant in USE AND USB ⇒ Combination score = 5		
	If conflicting features were present or if the features were not typical ⇒ combination score would be 3 or 4, depending on the most significant features		
Yoon et al. [40]	If USE and USB were different ⇒ Combination score would be the highest score of either of them	↔	↑
Sohn et al. [22]	If ES > BIRADS ⇒ Upgrade of BIRADS	↓	↑
	If ES < BIRADS ⇒ Downgrade of BIRADS		
	The degree of grading modification was based on each examiners' experience and confidence with USE.		
Zhi et al. [27]	Not defined	↑	↑
Current study (2011)	Conjunctive combination	↓	↑
	If USE AND USB is malignant ⇒ combination is malignant		
	If USE OR USB are benign ⇒ combination is benign		
	Disjunctive combination	↑	↓
	If USE OR USB is malignant ⇒ combination is malignant		
	If USE AND USB are benign ⇒ combination is benign		

USE ultrasound elastography, *USB* B-mode ultrasound, *ES* elasticity score, *BIRADS* Breast Imaging-Reporting and Data System

References

1. Garra BS (2007) Imaging and estimation of tissue elasticity by ultrasound. Ultrasound Q 23:255–268
2. Hatzung G, Grunwald S, Zygmunt M et al (2010) Sonoelastography in the Diagnosis of Malignant and Benign Breast Lesions: Initial Clinical Experiences. Ultraschall Med 31:596–603
3. Itoh A, Ueno E, Tohno E et al (2006) Breast disease: clinical application of US elastography for diagnosis. Radiology 239:341–350
4. Wells PN, Liang HD (2011) Medical ultrasound: imaging of soft tissue strain and elasticity. J R Soc Interface. [Epub ahead of print]
5. Regini E, Bagnera S, Tota D et al (2010) Role of sonoelastography in characterising breast nodules. Preliminary experience with 120 lesions. Radiol Med 115:551–562
6. Chang JM, Moon WK, Cho N, Kim SJ (2011) Breast mass evaluation: factors influencing the quality of US elastography. Radiology 259:59–64
7. Navarro B, Ubeda B, Vallespi M et al (2011) Role of elastography in the assessment of breast lesions: preliminary results. J Ultrasound Med 30:313–321
8. Lee JH, Kim SH, Kang BJ et al (2011) Role and clinical usefulness of elastography in small breast masses. Acad Radiol 18:74–80
9. Fischer T, Peisker U, Fiedor S et al. (2011) Significant Differentiation of Focal Breast Lesions: Raw Data-Based Calculation of Strain Ratio. Ultraschall Med. [Epub ahead of print]
10. Yerli H, Yilmaz T, Kaskati T, Gulay H (2011) Qualitative and semiquantitative evaluations of solid breast lesions by sonoelastography. J Ultrasound Med 30:179–186
11. Kumm TR, Szabunio MM (2010) Elastography for the characterization of breast lesions: initial clinical experience. Cancer Control 17:156–161
12. Thomas A, Degenhardt F, Farrokh A et al (2010) Significant differentiation of focal breast lesions: calculation of strain ratio in breast sonoelastography. Acad Radiol 17:558–563
13. Thomas A, Fischer T, Frey H et al (2006) Real-time elastography–an advanced method of ultrasound: First results in 108 patients with breast lesions. Ultrasound Obstet Gynecol 28:335–340
14. Thomas A, Kummel S, Fritzsche F et al (2006) Real-time sonoelastography performed in addition to B-mode ultrasound and

mammography: improved differentiation of breast lesions? Acad Radiol 13:1496–1504

15. Thomas A, Warm M, Hoopmann M et al (2007) Tissue Doppler and strain imaging for evaluating tissue elasticity of breast lesions. Acad Radiol 14:522–529

16. Wojcinski S, Farrokh A, Weber S et al (2010) Multicenter Study of Ultrasound Real-Time Tissue Elastography in 779 Cases for the Assessment of Breast Lesions: Improved Diagnostic Performance by Combining the BI-RADS(R)-US Classification System with Sonoelastography. Ultraschall Med 31:484–491

17. Parajuly SS, Lan PY, Yan L et al (2010) Breast elastography: a hospital-based preliminary study in China. Asian Pac J Cancer Prev 11:809–814

18. Ciurea AI, Bolboaca SD, Ciortea CA et al (2011) The influence of technical factors on sonoelastographic assessment of solid breast nodules. Ultraschall Med 32(Suppl 1):S27–S34

19. Chung SY, Moon WK, Choi JW et al (2010) Differentiation of benign from malignant nonpalpable breast masses: a comparison of computer-assisted quantification and visual assessment of lesion stiffness with the use of sonographic elastography. Acta Radiol 51:9–14

20. Raza S, Odulate A, Ong EM et al (2010) Using real-time tissue elastography for breast lesion evaluation: our initial experience. J Ultrasound Med 29:551–563

21. Zhi H, Xiao XY, Yang HY et al (2010) Ultrasonic elastography in breast cancer diagnosis: strain ratio vs. 5-point scale. Acad Radiol 17:1227–1233

22. Sohn YM, Kim MJ, Kim EK et al (2009) Sonographic elastography combined with conventional sonography: how much is it helpful for diagnostic performance? J Ultrasound Med 28:413–420

23. Schaefer FK, Heer I, Schaefer PJ et al (2009) Breast ultrasound elastography—Results of 193 breast lesions in a prospective study with histopathologic correlation. Eur J Radiol 77:450–456

24. Zhu QL, Jiang YX, Liu JB et al (2008) Real-time ultrasound elastography: its potential role in assessment of breast lesions. Ultrasound Med Biol 34:1232–1238

25. Vanhoutte A, Fellah L, Galant C et al (2008) Contribution of sonoelastography to the characterization of breast lesions. JBR-BTR 91:187–194

26. Tan SM, Teh HS, Mancer JF, Poh WT (2008) Improving B mode ultrasound evaluation of breast lesions with real-time ultrasound elastography—a clinical approach. Breast 17:252–257

27. Zhi H, Ou B, Luo BM et al (2007) Comparison of ultrasound elastography, mammography, and sonography in the diagnosis of solid breast lesions. J Ultrasound Med 26:807–815

28. Scaperrotta G, Ferranti C, Costa C et al (2008) Role of sono-elastography in non-palpable breast lesions. Eur Radiol 18:2381–2389

29. Cho N, Moon WK, Park JS et al (2008) Nonpalpable breast masses: evaluation by US elastography. Korean J Radiol 9:111–118

30. Satake H, Nishio A, Ikeda M et al (2011) Predictive value for malignancy of suspicious breast masses of BI-RADS categories 4 and 5 using ultrasound elastography and MR diffusion-weighted imaging. AJR Am J Roentgenol 196:202–209

31. Tardivon A, El Khoury C, Thibault F et al (2007) Elastography of the breast: a prospective study of 122 lesions. J Radiol 88:657–662

32. Giuseppetti GM, Martegani A, Di Cioccio B, Baldassarre S (2005) Elastosonography in the diagnosis of the nodular breast lesions: preliminary report. Radiol Med 110:69–76

33. Kadour MJ, Adams R, English R et al (2010) Slip imaging: reducing ambiguity in breast lesion assessment. Ultrasound Med Biol 36:2027–2035

34. Bartolotta TV, Ienzi R, Cirino A et al (2011) Characterisation of indeterminate focal breast lesions on grey-scale ultrasound: role of ultrasound elastography. Radiol Med 116:1027–1038

35. Zhuang C, Xiao Y (2009) Ultrasonic elastography and molybdenum X-ray photography in differential diagnosis of breast diseases. Zhong Nan Da Xue Xue Bao Yi Xue Ban 34:67–71

36. Zhang XF, Liu XM, Bao XF et al (2006) Application of real-time tissue elastography in diagnosis of breast cancer. Zhejiang Da Xue Xue Bao Yi Xue Ban 35:444–447

37. Gong X, Xu Q, Xu Z et al (2011) Real-time elastography for the differentiation of benign and malignant breast lesions: a meta-analysis. Breast Cancer Res Treat 130:11–18

38. Sadigh G, Carlos RC, Neal CH, Dwamena BA (2011) Accuracy of ultrasound elastography for differentiation of malignant and benign breast abnormalities: a meta-analysis. Radiology (Accepted for publication)

39. Tohno E, Ueno E (2008) Current improvements in breast ultrasound, with a special focus on elastography. Breast Cancer 15:200–204

40. Yoon JH, Kim MH, Kim EK et al (2011) Interobserver variability of ultrasound elastography: how it affects the diagnosis of breast lesions. AJR Am J Roentgenol 196:730–736

41. Cochrane Diagnostic Test Accuracy Review Working Group. http://srdta.cochrane.org/handbook-dta-reviews. Accessed 15 Aug 2011

42. Ament A, Hasman A (1993) Optimal test strategy in the case of two tests and one disease. Int J Biomed Comput 33:179–197

43. American College of Radiology (2009) Breast imaging reporting and data system (BI-RADS®) ultrasound. http://www.acr.org/SecondaryMainMenuCategories/quality_safety/BIRADSAtlas.aspx. Accessed 12 Oct 2010

44. Whiting P, Rutjes AW, Reitsma JB et al (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3:25

45. Chu H, Cole SR (2006) Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. J Clin Epidemiol 59:1331–1332 (author reply 1332–1333)

46. Arends LR, Hamza TH, van Houwelingen JC et al (2008) Bivariate random effects meta-analysis of ROC curves. Med Decis Mak 28:621–638

47. Cronin P, Dwamena BA, Kelly AM et al (2008) Solitary pulmonary nodules and masses: a meta-analysis of the diagnostic utility of alternative imaging tests. Eur Radiol 18:1840–1856

48. Chappell FM, Raab GM, Wardlaw JM (2009) When are summary ROC curves appropriate for diagnostic meta-analyses? Stat Med 28:2653–2668

49. Deeks JJ, Macaskill P, Irwig L (2005) The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol 58:882–893

50. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327:557–560

51. Dwamena, BA (2007) MIDAS: Stata module for meta-analytical integration of diagnostic test accuracy studies. Statistical Software Components S456880. Boston College Department of Economics, Boston. http://ideas.repec.org/c/boc/bocode/s456880.html

52. Fleury Ede F, Fleury JC, Oliveira VM et al (2009) Proposal for the systematization of the elastographic study of mammary lesions through ultrasound scan. Rev Assoc Med Bras 55:192–196

53. Fleury Ede F, Fleury JC, Piato S, Roveda D Jr (2009) New elastographic classification of breast lesions during and after compression. Diagn Interv Radiol 15:96–103

54. Wang Z, Yang T, Wu Z et al (2010) Correlation between elastography score and strain rate ratio in breast small tumor. Zhong Nan Da Xue Xue Bao Yi Xue Ban 35:928–932

55. Jones CM, Athanasiou T (2009) Diagnostic accuracy meta-analysis: review of an important tool in radiological research and decision making. Br J Radiol 82:441–446
56. Metz CE (1978) Basic principles of ROC analysis. Semin Nucl Med 8:283–298
57. Moon HJ, Kim MJ, Kwak JY, Kim EK (2010) Probably benign breast lesions on ultrasonography: a retrospective review of ultrasonographic features and clinical factors affecting the BI-RADS categorization. Acta Radiol 51:375–382
58. Gutwein LG, Ang DN, Liu H et al (2011) Utilization of minimally invasive breast biopsy for the evaluation of suspicious breast lesions. Am J Surg 202:127–132
59. Bennett ML, Welman CJ, Celliers LM (2011) How reassuring is a normal breast ultrasound in assessment of a screen-detected mammographic abnormality? A review of interval cancers after assessment that included ultrasound evaluation. Clin Radiol 66:928–939