RESEARCH ARTICLE

Diagnostic Accuracy of Ultrasonograph Guided Fine-needle Aspiration Cytologic in Staging of Axillary Lymph Node Metastasis in Breast Cancer Patients: a Meta-analysis

Xi-Wen Wang[1&], Yun-Hui Xiong[2&], Xiao-Qing Zen[1], Hai-Bo Lin[1], Qing-Yi Liu[1*]

Abstract

Purpose: To evaluate the diagnostic accuracy of ultrasonograph and fine-needle aspiration cytologic examination (USG-FNAC) in the staging of axillary lymph node metastasis in breast cancer patients. Methods: We conducted an electronic search of the literature addressing the performance of USG-FNAC in diagnosis of axillary lymph node metastasis in databases such as Pubmed, Medline, Embase, Ovid and Cochrane library. We introduced a series of diagnostic test indices to evaluate the performance of USG-FNAC by the random effect model (REM), including sensitivity, specificity, likelihood ratios, and diagnostic odds ratios and area under the curve (AUC). Results: A total of 20 studies including 1371 cases and 1289 controls were identified. The pooled sensitivity was determined to be 0.66 (95% CI 0.64-0.69), specificity 0.98 (95% CI 0.98-0.99), positive likelihood ratio 22.7 (95% CI 15.0-34.49), negative likelihood ratio 0.32 (95% CI 0.25-0.41), diagnostic OR 84.2 (95% CI 53.3-133.0). Due to the marginal threshold effect found in some indices of diagnostic validity, we used a summary SROC curve to aggregate data, and obtained a symmetrical curve with an AUC of 0.942. Conclusion: The results of this meta-analysis indicated that the USG-FNAC techniques have acceptable diagnostic validity indices and can be used for early staging of axillary lymph node in breast cancer patients.

Keywords: Breast cancer - USG-FNAC - meta-analysis - sensitivity - specificity

Asian Pacific J Cancer Prev, **13 (11)**, 5517-5523

Introduction

In the developed countries, breast cancer is one of the most common modalities of malignancy in women and the global incidence of breast cancer is still increasing (Parkin et al., 2005). At present, screening for breast cancer in women is regarded as effective way to reduce incidence of breast cancer. So a rapid and accurate diagnosis tool is crucial and fundamental in the implementation of breast cancer screening. The diagnosis of axillary staging, however, is an utmost question worthy of careful consideration in the management of breast cancer (Fisher et al., 1993; Samphao et al., 2008). In breast cancer patients the status of axillary lymph node metastasis is important prognostic indicator and determinant for selecting patients who should receive adjuvant treatment (Fisher et al., 1983). Consequently, the diagnosis of the axillary lymph node metastasis status possesses great priority in the treatment of breast cancer. To date, some techniques, such as clinical examination, ultrasonograph, computed tomography, magnetic resonance imaging, sentinel lymph node dissection, fine-needle aspiration biopsy, axillary lymph node dissection (ALND), et al were adopted to explore the lymph node status of a breast cancer patient. But the performance of any technique alone is not perfect, especially disappointed in the detection of impalpable node. Also the accuracy, safety and cost varied greatly among them. Clinical examination alone is neither a sensitive nor reliable way to ascertain lymph node status, because metastatic lymph nodes are often impalpable and reactive lymph nodes may be mistaken for metastasis (Sacre , 1983; Pamilo et al., 1989; De Freitas et al., 1991). The accuracy of preoperative ultrasonographic diagnosis of nodal metastasis has improved with the development of high-frequency ultrasonography technology. Axillary ultrasonography is increasingly being used to improve the staging of breast cancer patients who have negative axillary lymph nodes on physical examination (Herrada et al., 1997; de Kanter et al., 1999). But due to overlapping sonographic features of benign/reactive and suspicious/ metastatic lymph nodes, the value of ultrasonography in the diagnosis of lymph nodes metastasiss had been discounted.

For many years, the lymph node status of a breast cancer patient was determined by performing an axillary lymph node dissection (ALND). However, the

*[1]Bao'an Maternal and Child Health Hospital, [2]Fuyong People's Hospital, Shenzhen, China [&]Equal contributors *For correspondence: liuqingyi_2003@163.com*

disadvantage of this method is the significant morbidity that is associated with it, e.g. lymph oedema of the arm with a decreased ability for movement (Lotze et al., 1981; Kissin et al., 1986; Shaw et al., 1990; vens et al., 1992). SLN biopsy was proved to stage accurately the axillary LN status in breast cancer patients and has been adopted worldwide as an alternative to the ALND (Schwartz et al., 2002). However, the main weakness to SLNB are that the technique itself is quite complex and that the adequacy of the biopsy relies greatly on the skill and experience of the surgeon. CT scans and MRI are limited to diagnosing enlargement of lymph nodes without being able to differentiate between those infiltrated with cancer from hyperplastic glands. Other drawbacks of CT and MRI include the high cost and the difficulty of obtaining material for pathologic analysis. New techniques such as positron emission tomography are even costlier than CT and MRI, and experiences with this method are limited (Adler et al., 1993). The combination of different methods to axillary lymph nodes staging was reported in published literatures and significantly improved the accuracy of diagnosis in axillary lymph nodes metastasis. The most common modality is the combination of ultrasonography with FNAC. FNAC of nonpalpable axillary lymph nodes can improve markedly the specificity of both physical examination and US alone in detecting metastatic lymph nodes (Krishnamurthy et al., 2002). In conclusion, US-guided FNAC of nonpalpable indeterminate and suspicious axillary lymph nodes is a simple, minimally invasive and reliable technique for the initial determination of axillary lymph node status in breast carcinoma patients (Krishnamurthy et al., 2002).

Up to now, many studies have reported the performance of USG-FNAC in the staging of axillary lymph node. But the sensitivity and specificity in the diagnosis of axillary lymph node metastasis diversified greatly. In order to assess the performance of ultrasound-guided fine-needle aspiration cytology objectively and find out a optimal procedure for staging of axillary lymph nodes, we performed the meta-analysis.

Materials and Methods

Literature search

We conducted an electronic search for literatures addressing the performance of US-guide FNAC in diagnosis of the axillary lymph node metastasis in databases such as Pubmed, Medline, Embase, Ovid, Cochrane library. The search strategy: ("breast cancer" [ti] OR "breast carcinoma"[ti]) AND (FNAC OR "fine-needle aspiration cytology" OR "fine needle aspiration cytology") AND sensitivity AND specificity AND (US OR ultrasound OR ultrasonograph) was used to collecte related articles up to July 2012.

Selection of articles

An eligible article must meet the following inclusion criteria: (1) US-guide FNAC method was adopted to stage axillary lymph node in study. (2) true positive (TP), false positive (FP), false negative (FN), true negative (TN) was reported , or can be calculated. (3) Reference standard refer to histopathologic analysis. (4) axillary lymph node was impalpable.

Quality assessment of studies included

Quality assessment of studies included was achieved by use of the quality assessment of diagnostic studies (QUADAS) instrument, a quality assessment tool specifically developed for systematic reviews of diagnostic accuracy studies (Whiting et al., 2003; Whiting et al., 2006).

Data extraction

The specialized form for data extraction was tabulated according to the requirement of meta-analysis. The data needed to be collected included: First author, publication year, origin, design, reference standard, patient selection, blind design, number of patients, QUADAS score, number of cases, TP (true positive), FP (false positive), TN (true negative), FN (false negative). Sensitivity, specificity and accurate can be calculated by the combination of TP, FP, TN and FN. Two authors reviewed and extracted the needed data independently from eligible articles. When disagreement appeared, the third author was consulted and problem was revolved by the major of vote.

Statistical analysis

Heterogeneity analysis: A common source of heterogeneity among diagnostic tests rooted from threshold effect, defined as"the use of different criteria in each study to determine whether the test is positive or negative". Representation of accuracy estimates from each study in a receiver operating characteristic (ROC) space and computation of Spearman correlation coefficient between the log (SEN) and log (1-SPE) were assessed for threshold effect. A typical pattern of "shoulder arm" plot in a ROC space and a strong positive correlation would suggest threshold effect (Moses et al., 1993; Zamora et al., 2006). To assess between-study heterogeneity (other than threshold effect) and between-study inconsistency Cochran Q statistic and inconsistency index (I^2) were calculated and the level of significance for the corresponding P-value was set at $P = 0.10$. Heterogeneity was considered low if the I^2 value was 25% or less, moderate if the value was between 25% and 50%, high if between 50% and 75% and very high if greater than 75% (Higgins et al., 2003).

Data synthesis

The data synthesis for the accuracy of USG-FNAC was made by calculating pooled estimates of sensitivity, specificity, likelihood ratios, and diagnostic odds ratios. The estimates can be performed by the fixed effect model (FEM) or by the random effect model (REM) to incorporate variation among studies, and the output can be presented graphically as forest plots. Due to anticipated inter-study heterogeneity, a random effects analysis model (DerSimonian Laird) (DerSimonian et al., 1986) was applied in all meta-analytic calculations because it provides more conservative estimates of the pooled data. If heterogeneity among studies is present, the accuracy data can be pooled by fitting a summary ROC (sROC) curve and summarizing that curve by means of the area

Table 1. General Characteristics of all 21 Eligible Studies Included in Meta-analysis

First author, year	Origin	Design	Reference standard	Patient selection	Blind design	N of patients	QUADAS score
Bonnema J, 1997	Netherlands	prospective	histological	consecutive	NA	81	10
Krishnamurthy S, 2002	USA	retrospective	histological	consecutive	NA	103	11
Kuenen-Boumeester V, 2003	Netherlands	prospective	histological	consecutive	NA	134	12
Sapino A, 2003	Italy	prospective	histological	consecutive	NA	85	9
Deurloo EE, 2003	Netherlands	prospective	histological	consecutive	NA	66	11
Bedrosian I, 2003	UAS	retrospective	histological	consecutive	NA	22	12
van Rijk MC, 2006	Netherlands	prospective	histological	consecutive	NA	176	10
Duchesne N, 2005	Canada	retrospective	histological	consecutive	NA	40	11
Popli MB, 2006	India	prospective	histological	consecutive	NA	24	12
Ciatto S, 2007	Italy	retrospective	histological	consecutive	NA	435	10
Altomare V, 2007	Italy	retrospective	histological	consecutive	NA	70	11
Alkuwar E, 2008	Canada	retrospective	histological	consecutive	NA	115	13
Tahir M, 2008	UK	prospective	histological	consecutive	NA	38	10
Luparia A, 2010	Italy	retrospective	histological	consecutive	NA	129	12
Baruah BP, 2010	UK	retrospective	histological	consecutive	NA	502	11
Schiettecatte A, 2011	Belgium	retrospective	histological	consecutive	NA	148	11
Jung J, 2010	Korea	retrospective	histological	consecutive	NA	74	12
Hayes BD, 2011	Ireland	retrospective	histological	consecutive	NA	161	11
Carroll PA, 2011	Ireland	retrospective	histological	consecutive	NA	188	10
Devaraj S, 2011	UK	prospective	histological	consecutive	NA	128	12
Podkrajsek M, 2005	Slovenia	retrospective	histological	consecutive	NA	44	11

NA, not available; QUADAS, quality assessment of diagnostic studies

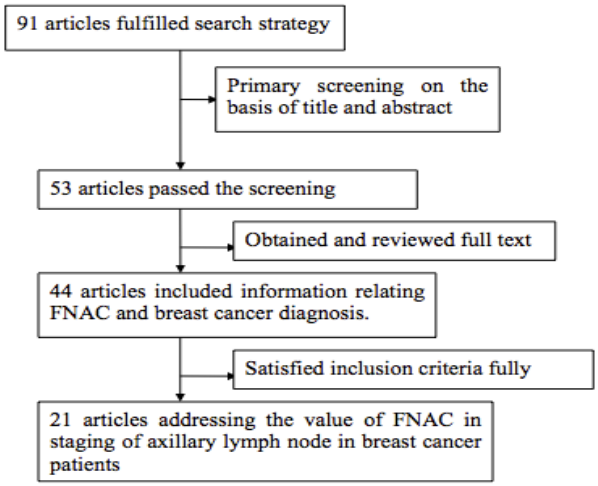

Figure 1. Flow Chart of Literature Search and Selection

under the curve (AUC).

Publication bias

An estimate of potential publication bias was obtained using a Begg funnel plot, in which the standard error of log (DOR) for each study was plotted against its log (OR). An asymmetric plot suggests potential publication bias. Funnel plot asymmetry was assessed by Egger's linear regression test. Significant publication bias is confirmed when the P-value for bias is <0.05 in Egger's test.

Sensitivity Analyses

The pooled estimates were reappraised when suspicious studies were excluded, and the reappraised results were compared with the original results to assess stability and reliability of our meta-analysis.

Statistic software package

The homogeneity test, threshold effect analysis, pooled weighted sensitivity and specificity, sROC curve and sensitivity analysis were performed by using Meta-Disc version 1.4 (Zamora et al., 2006).

Results

Literature search and selection

After careful computer search and manual search, a total of 91 articles were captured. Upon primary screening by title and abstract, 38 articles were abandoned because of irrelevance to FNAC clinical practice in diagnosis of breast cancer. Afterwards, the full text of the remaining 53 articles were obtained and reviewed. Consequently, 44 articles included information relating FNAC and breast cancer diagnosis. But only 21 of 44 articles satisfied the requirements of the meta-analysis fully (Bonnema et al., 1997; Krishnamurthy et al., 2002; Bedrosian et al., 2003; Deurloo et al., 2003; Kuenen-Boumeester et al., 2003; Sapino et al., 2003; Duchesne et al., 2005; Podkrajsek et al., 2005; Popli et al., 2006; van Rijk et al., 2006; Altomare et al., 2007; Ciatto et al., 2007; Alkuwari et al., 2008; Tahir et al., 2008; Baruah et al., 2010; Jung et al., 2010; Luparia et al., 2010; Carroll et al., 2011; Devaraj et al., 2011; Hayes et al., 2011; Schiettecatte et al., 2011), addressing the value of FNAC in staging of axillary lymph node in breast cancer patients. The detailed procedure was seen in Figure 1.

General characteristic of studies included

General characteristics of all 21 eligible studies were listed in Table 1 (Bonnema et al., 1997; Krishnamurthy et al., 2002; Bedrosian et al., 2003; Deurloo et al., 2003; Kuenen-Boumeester et al., 2003; Sapino et al., 2003; Duchesne et al., 2005; Podkrajsek et al., 2005; Popli et al., 2006; van Rijk et al., 2006; Altomare et al., 2007; Ciatto et al., 2007; Alkuwari et al., 2008; Tahir et al., 2008; Baruah et al., 2010; Jung et al., 2010; Luparia et al.,

Table 2. Main Data for Diagnostic Accuracy of the Studies Included in the Meta-analysis

First author, year	N of cases	TP	FP	FN	TN	Sensitivity (%)	Specificity (%)	PPV(%)	NPV(%)	Accuracy (%)
Bonnema J, 1997	81	39	0	10	32	79.59	100	100	76.19	87.65
Krishnamurthy S, 2002	103	51	16	12	24	80.95	60	76.12	66.67	72.82
Kuenen-Boumeester V, 2003	134	37	3	28	66	56.92	95.65	92.5	70.21	76.87
Sapino A, 2003	85	49	0	6	30	89.09	100	100	83.33	92.94
Deurloo EE, 2003	66	37	0	12	17	75.51	100	100	58.62	81.82
Bedrosian I, 2003	22	3	0	9	10	25	100	100	52.63	59.09
van Rijk MC, 2006	176	58	1	36	81	61.7	98.8	98.3	69.2	79
Duchesne N, 2005	40	29	1	2	8	93.55	88.89	96.67	80	92.5
Popli MB, 2006	24	15	0	4	5	78.9	100	100	55.6	83.3
Ciatto S, 2007	435	199	7	75	154	72.6	95.7.0	96.6	67.2	81.5
Altomare V, 2007	70	30	0	14	26	68.18	100	100	65	80
Alkuwar E, 2008	115	49	0	26	40	65.3	100	100	60.6	77.39
Tahir M, 2008	38	8	0	9	21	47.06	100	100	70	76.32
Luparia A, 2010	129	71	0	9	49	88.75	100	100	84.48	93.02
Baruah BP, 2010	502	39	0	98	365	28.47	100	100	78.83	80.48
Schiettecatte A, 2011	148	34	0	34	80	50	100	100	70.18	77.03
Jung J, 2010	74	32	1	6	35	84.21	97.22	97	85.37	90.54
Hayes BD, 2011	161	57	1	29	74	66.28	98.67	98.28	71.84	81.37
Carroll PA, 2011	188	59	0	34	95	63.44	100	100	73.64	81.91
Devaraj S, 2011	128	30	0	19	79	61.22	100	100	80.61	85.16
Podkrajsek M, 2005	44	32	1	4	7	88.89	87.5	96.97	63.64	88.64

TP, true positive; FP, false positive; FN, false negative; TN, true negative; PPV, Positive Predictive Value; NPV, Negative Predictive Value

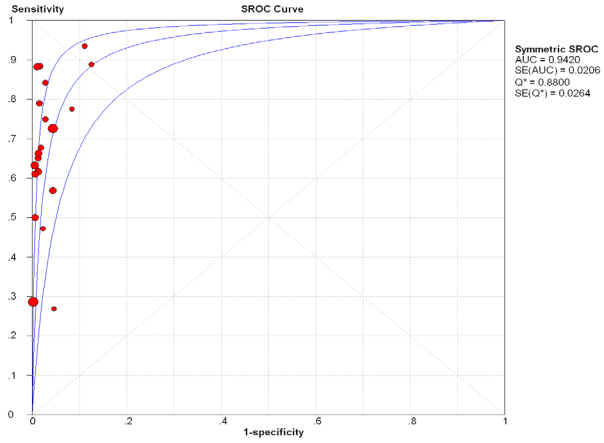

Figure 2. SROC Curve for the USG-FNAC Meta-analysis. Each red circle represents an individual research study in the meta-analysis, with the size of the diamond directly proportional to the sample size of the study. The best fit curve (middle curve) lies between the other 2 curves, which demarcate its 95% CI. The diamond denotes the Q*-point

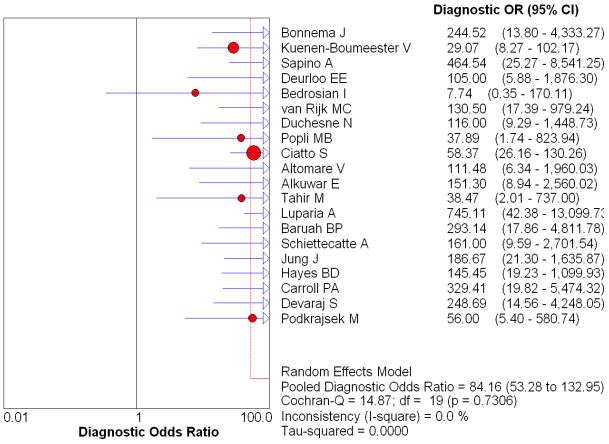

Figure 3. Forrest Plot of the Diagnotic Odds Ratio of Each Individual Study, Pooled Diagnotic Odds Ratio, and I^2 Statistic for Heterogeneity

2010; Carroll et al., 2011; Devaraj et al., 2011; Hayes et al., 2011; Schiettecatte et al., 2011), including first author, publication year, origin, design, reference standard, patient selection, blind design, number of patients and QUADAS score.

Main data collected for meta-analysis

Main data for meta-analysis included number of cases, TP (true positive), FP (false positive), TN (true negative), FN (false negative). Sensitivity, specificity, PPV, NPV and accurate were calculated by the combination of TP, FP, TN and FN. These data can be seen in Table 2.

Heterogeneity analysis

Given that individual study may have great influence on overall results, sensitivity analysis was performed firstly. Unfortunately, a study author by Krishnamurthy S was picked out and changed pooled result dramatically (Krishnamurthy et al., 2006). Consequently the study was ruled out before heterogeneity analysis.

A atypical pattern of "shoulder arm" plot in a ROC space and a mild positive correlation (Spearman correlation coefficient: ρ=0.412, P=0.071) suggested no obvious threshold effect (Figure 2). There was marked heterogeneity in the studies in terms of sensitivity (chi-square =184.00; P<0.05; I^2 =89.7%), specificity (chi-square =38.29; P<0.05; I^2 =50.4%), LR- (Cochran-Q=209.91; P<0.05; I^2 =90.9%).However, no significant heterogeneity present in LR+ (Cochran-Q=16.47; P>0.05; I^2 =0.0%) and diagnostic OR (Cochran-Q=14.87; P>0.05; I^2 =0.0%).

Synthesis of the data

Because of significant heterogeneity, the diagnostic indices were calculated using the random effects model. Using a forest plot, pooled sensitivity was determined to be 0.66 (95% CI 0.64–0.69), specificity 0.98 (95%

CI 0.98–0.99). The positive likelihood ratio was 22.73 (95% CI 14.99–34.49), and the negative likelihood ratio was 0.32 (95% CI 0.25–0.41). The pooled diagnostic OR was 84.17 (95% CI 53.28–132.95) (Figure 3).Due to the marginal threshold effect found in some indices of diagnostic validity, we used a summary sROC curve to aggregate data, and obtained a symmetrical curve with an AUC of 0.942 that represented the technique's diagnostic performance. The Q* value 0.880 shows the test's optimum cutoff point, where sensitivity and specificity reach their maximum value; it is an overall measure of the technique's accuracy.

Publication bias

The inspection of funnel plots and the statistical tests for publication bias revealed an obvious effect of publication bias (In egger test, P<0.05).

Sensitivity analysis

Sensitivity analyses were carried out by limiting a single study at a time into the meta-analysis. When the study author by Krishnamurthy S was dropped (Krishnamurthy et al., 2002), the pooled DOR changed markedly. This suggested that the study influence greatly the overall result.

Discussion

This meta-analysis synthesizes the current knowledge about early staging of axillary lymph node by the technique of USG-FNAC in breast cancer patients. In this meta-analysis, we calculate an overall sensitivity of 0.66 (95% CI 0.64–0.69), which means that the frequency of test FN was 34 per 100 diagnosed patients. Specificity was 0.98 (95% CI 0.98–0.99), which means that the frequency of FP of the technique (diagnosing patients who were in fact disease-free) was 2 per 100 patients from 20 studies fulfilling all inclusion and exclusion criteria. The pooled Diagnostic Odds Ratios was high to 84.16 (95% CI 53.28 to 132.95). The accuracy ranged from 0.59 to 0.93). The AUC of sROC reached up to 0.942. All index indicated that the accuracy of USG-FNAC was fairly satisfactory.

However, there is notable heterogeneity among individual studies. Therefore, it is critical to investigate the source of heterogeneity to determine the potential impact factors and to evaluate the appropriateness of statistical pooling of accuracy estimates from various studies. The heterogeneity in a meta-analysis can have diverse sources, including differences in the test procedures, differences in the subject populations, different inclusion criteria, differences in the study designs, different definitions to classify self reported symptoms, or a combination of these factors (Irwig et al., 1995). Other contributing factors may be differences in clinical diagnostic skills or biopsy method, and/or interpretation of histology results across studies.

Of course, significant heterogeneity concerning the results between studies, probably caused by a threshold effect. In the meta-analysis, however, threshold effect was marginal (Spearman correlation coefficient: 0.412, P= 0.071). Because of the trade-off nature between sensitivity and specificity, meta-analysis of diagnostic tests using these conventional expressions gives rise to statistical challenges. Simple synthesis of sensitivity and specificity may not be the most appropriate approach, as it ignores threshold differences. More recent evidence suggests that a more important index to evaluate the accuracy of a given test result as a diagnostic tool is the pooled DOR (Glas et al., 2003). The DOR offers considerable advantages in meta-analysis of diagnostic studies as it combines results from different studies into summary estimates with increased precision. The advantage of DOR was demonstrated perfectly in the meta-analysis. The inconsistency coefficient I^2 for DOR among included studies was so small that it draw near to zero (heterogeneity chi-squared = 14.87, P = 0.731).

Meta-analysis combines or integrates the results of several independent studies. The quality and reliability of a meta-analysis depends on the quality of included studies. We use the QUADAS tool for assessing methodological quality of individual studies. This tool was specifically developed for quality assessment of diagnostic accuracy studies included in systematic reviews and has been used to help identify severe methodological shortcomings (Whiting et al., 2006). Most included studies in this meta-analysis had a suboptimal design in regard to the reporting of selection criteria, the description of the execution of the reference standard, the interpretation of the reference standard results without knowledge of the index test results, the interpretation of the index test results without knowledge of the reference standard, reporting of uninterpretable and/or intermediate test results, or explanation of withdrawals from the study (Table 1).

Since publication biases would tend to exaggerate clinical effects resulting in potentially erroneous clinical decision making, it is important to assess the likely extent of the bias and its potential impact on the conclusions (Song et al., 2002). Publication bias can be visually examined after construction of a funnel plot and quantitatively detected by egger test. The inspection of funnel plots and the statistical tests for publication bias revealed an obvious effect of publication bias (In egger test, P<0.05). The potential reasons may be attributed to that studies with optimistic results may be published easier than studies with unfavorable results, and studies with large sample size may be published easier than studies with small sample size.

In order to evaluate the robustness of the study, sensitivity analyses were carried out by limiting a single study at a time into the meta-analysis. The pooled estimates were reappraised when suspicious studies were excluded, and the reappraised results were compared with the original results to assess stability and reliability of our meta-analysis. When the study author by Krishnamurthy S was dropped, the pooled DOR changed markedly. This suggested that the study influence greatly the overall result. So the study Krishnamurthy S was given up in the course of statistical synthesis.

In this meta-analysis, there were several shortcomings which must be acknowledged. First, the limited number of relevant research studies for the statistical analysis because of incomplete data reported, high variability in

the methodology, and heterogeneity of the data. Second, since yet unpublished studies, gray literature, and reports from commercial enterprises were excluded, potential publication bias was inevitable.Third, we only collected literatures written in English. Related studies written in other languages may be missed. Thus language bias appeared naturally. Fourth, review bias is also an issue worthy of attention. Review bias refers to a situation where persons interpreting the index test have knowledge of the reference standard or vice versa, when persons interpreting the reference standard have knowledge of the index test (Begg et al., 1987). In our meta-analysis, it was very unclear whether this did or did not occur because the majority of the studies did not report whether blinding during testing was done.

The results of this meta-analysis indicated that the USG-FNAC techniques have acceptable diagnostic validity indices and can be used for early staging of axillary lymph node in breast cancer patients.

Acknowledgements

The authors are fully responsible for all content and have not received any financial support related to the development of this article. The author(s) declare that they have no competing interests.

References

Adler LP, Crowe JP, Al-Kaisi NK, et al (1993). Evaluation of breast masses and axillary lymph nodes with [F-18] 2-deoxy-2-fluoro- D-glucose PET. *Radiology*, **187**, 743-50

Alkuwari E, Auger M (2008). Accuracy of fine-needle aspiration cytology of axillary lymph nodes in breast cancer patients: a study of 115 cases with cytologic-histologic correlation. *Cancer*, **114**, 89-93.

Altomare V, Guerriero G, Carino R, et al (2007). Axillary lymph node echo-guided fine-needle aspiration cytology enables breast cancer patients to avoid a sentinel lymph node biopsy. Preliminary experience and a review of the literature. *Surg Today*, **37**, 735-9.

Baruah BP, Goyal A, Young P, et al (2010). Axillary node staging by ultrasonography and fine-needle aspiration cytology in patients with breast cancer. *Br J Surg*, **97**, 680-3.

Bedrosian I, Bedi D, Kuerer HM, et al (2003). Impact of clinicopathological factors on sensitivity of axillary ultrasonography in the detection of axillary nodal metastases in patients with breast cancer. *Ann Surg Oncol*, **10**, 1025-30.

Begg CB (1987). Biases in the assessment of diagnostic tests. *Stat Med*, **6**, 411-23.

Bonnema J, van Geel AN, van Ooijen B, et al (1997) .Ultrasound-guided aspiration biopsy for detection of nonpalpable axillary node metastases in breast cancer patients: new diagnostic method. *World J Surg*, **21**, 270-4.

Carroll PA, O'Mahony D, McDermott R, et al (2011). Perioperative diagnosis of the positive axilla in breast cancer: a safe, time efficient algorithm. *Eur J Surg Oncol*, **37**, 205-10.

Ciatto S, Brancato B, Risso G, et al (2007). Accuracy of fine needle aspiration cytology (FNAC) of axillary lymph nodes as a triage test in breast cancer staging. *Breast Cancer Res Treat*, **103**, 85-91.

De Freitas R Jr, Costa MV, Schneider SV, et al (1991). Accuracy of ultrasound and clinical examination in the diagnosis of axillary lymph node metastasis in breast cancer. *Eur J Surg Oncol*, **17**, 240-4.

de Kanter AY, van Eijck CH, van Geel AN, et al (1999). Multicentre study of ultrasonographically guided axillary node biopsy in patients with breast cancer. *Br J Surg,* **86**, 1459-62

DerSimonian R, Laird N (1986). Meta-analysis in clinical trials. *Control Clin Trials*, **7**, 177-88.

Deurloo EE, Tanis PJ, Gilhuijs KG, et al (2003). Reduction in the number of sentinel lymph node procedures by preoperative ultrasonography of the axilla in breast cancer. *Eur J Cancer*, **39**, 1068-73.

Devaraj S, Iqbal M, Donnelly J, et al (2011).Axillary ultrasound in invasive breast cancer: experience of our surgeons. *Breast J*, **17**,191-5.

Duchesne N, Jaffey J, Florack P, et al (2005). Redefining ultrasound appearance criteria of positive axillary lymph nodes. *Can Assoc Radiol J*, **56**, 289-96.

Fisher B, Bauer M, Wickerham D, et al (1983). Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer: an NSABP update. *Cancer*, **52**, 1551-7

Fisher ER, Anderson S, Redmond C, et al (1993). Pathologic findings from the national surgical adjuvant breast project protocol B-06. 10- year pathologic and clinical prognostic discriminants. *Cancer*, **71**, 2507-24.

Glas AS, Lijmer JG, Prins MH, et al (2003). The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol*, **56**, 1129-35.

Hayes BD, Feeley L, Quinn CM, et al (2011). Axillary fine needle aspiration cytology for pre-operative staging of patients with screen-detected invasive breast carcinoma. *J Clin Pathol,* **64**, 338-42.

Herrada J, Iyer RB, Atkinson EN, et al (1997). Relative value of physical examination, mammography, and breast sonography in evaluating the size of the primary tumor and regional lymph node metastasiss in women receiving neoadjuvant chemotherapy for locally advanced breast carcinoma. *Clin Cancer Res*, **3**, 1565-9.

Higgins J, Thompson S, Deeks J, et al (2003) .Measuring inconsistency in meta-analyses. *Br Med J*, **327**, 557-60.

Irwig L, Macaskill P, Glasziou P, et al (1995). Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol*, **48**, 119-30.

Jung J, Park H, Park J, et al (2010). Accuracy of preoperative ultrasound and ultrasound-guided fine needle aspiration cytology for axillary staging in breast cancer. *ANZ J Surg*, **80**, 271-5.

Kissin MW, Querci della Rovere G, Easton D, et al (1986). Risk of lymphoedema following the treatment of breast cancer. *Br J Surg,* **73**, 580-4.

Krishnamurthy S, Sneige N, Bedi DG, et al (2002). Role of ultrasound-guided fine-needle aspiration of indeterminate and suspicious axillary lymph nodes in the initial staging of breast carcinoma. *Cancer*, **95**, 982-8.

Krishnamurthy S, Sneige N, Bedi DG, et al (2002). Role of ultrasound-guided fine-needle aspiration of indeterminate and suspicious axillary lymph nodes in the initial staging of breast carcinoma. *Cancer*, **95**, 982-8.

Kuenen-Boumeester V, Menke-Pluymers M, de Kanter AY, et al (2003). Ultrasound-guided fine needle aspiration cytology of axillary lymph nodes in breast cancer patients. A preoperative staging procedure. *Eur J Cancer*, **39**, 170-4.

Lotze MT, Duncan MA (1981). Gerber LH, Woltering EA, Rosenberg SA. Early versus delayed shoulder motion following axillary dissection: a randomized prospective study. *Ann Surg*, **193**, 288-95.

Luparia A, Campanino P, Cotti R, et al (2010). Role of axillary

ultrasound in the preoperative diagnosis of lymph node metastases in patients affected by breast carcinoma. *Radiol Med*, **115**, 225-37.

Moses LE, Shapiro D, Littenberg B (1993). Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med*, **12**, 1293-316.

Pamilo M, Soiva M, Lavast EM (1989). Real time ultrasound, axillary mammography and clinical examination in the detection of axillary lymph node metastasiss in breast cancer patients. *J Ultrasound Med,* **8**, 115-20.

Parkin DM, Bray F, Ferlay J, et al (2005). Global cancer statistics, 2002. *Ca Cancer J Clin*, **55**, 74-108.

Podkrajsek M, Music MM, Kadivec M, et al (2005). Role of ultrasound in the preoperative staging of patients with breast cancer. *Eur Radiol,* **15**, 1044-50.

Popli MB, Sahoo M, Mehrotra N, et al (2006). Preoperative ultrasound-guided fine-needle aspiration cytology for axillary staging in breast carcinoma. *Australas Radiol*, **50**, 122-6.

Sacre RA (1986). Clinical evaluation of axillary lymph nodes compared to surgical and pathological findings. *Eur J Surg Oncol*, **12**, 169-73.

Samphao S, Eremin JM, El-Sheemy M, et al (2008). Management of the axilla in women with breast cancer: current clinical practice and a new selective targeted approach. *Ann Surg Oncol,* **15**, 1282-96.

Sapino A, Cassoni P, Zanon E, et al (2003). Ultrasonographically-guided fine-needle aspiration of axillary lymph nodes: role in breast cancer management. *Br J Cancer*, **88**, 702-6.

Schiettecatte A, Bourgain C, Breucq C, et al (2011). Initial axillary staging of breast cancer using ultrasound-guided fine needle aspiration: a liquid-based cytology study. *Cytopathology*, **22**, 30-5.

Schwartz GF, Giuliano AE, Veronesi U (2002). Proceedings of the consensus conference on the role of sentinel lymph node biopsy in carcinoma of the breast, April 19–22, 2001, Philadelphia, Pennsylvania. *Cancer*, **94**, 2542-51.

Shaw JH, Rumball EM (1990). Complications and local recurrence following lymphadenectomy. *Br J Surg*, **77**, 760-4.

Song F, Khan KS, Dinnes J, et al (2002). Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol*, **31**, 88-95.

Tahir M, Osman KA, Shabbir J, et al (2008). Preoperative axillary staging in breast cancer-saving time and resources. *Breast J*, **14**, 369-71.

van Rijk MC, Deurloo EE, Nieweg OE, et al (2006). Ultrasonography and fine-needle aspiration cytology can spare breast cancer patients unnecessary sentinel lymph node biopsy. *Ann Surg Oncol*, **13**, 31-5.

vens D, Hoe AL, Podd TJ, et al (1992). Assessment of morbidity from complete axillary dissection. *Br J Cancer*, **66**, 136-8.

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.

Whiting P, Weswood ME, Rutjes AW, et al (2006). Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies (update). *BMC Med Res Methodol*, **6**, 9.

Zamora J, Abraira V, Muriel A, et al (2006). Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol*, **6**, 31.