

Diffusion-weighted magnetic resonance imaging for the detection of metastatic lymph nodes in patients with lung cancer: A meta-analysis

GUANG-XIANG CHEN¹, MAO-HUA WANG², TING ZHENG¹,
GUANG-CAI TANG¹, FU-GANG HAN¹ and GUO-JIAN TU¹

Departments of ¹Radiology and ²Anesthesiology, The Affiliated Hospital of
Southwest Medical University, Luzhou, Sichuan 646000, P.R. China

Received June 14, 2016; Accepted December 6, 2016

DOI: 10.3892/mco.2017.1153

Abstract. The aim of the present meta-analysis was to evaluate the diagnostic value of diffusion-weighted imaging (DWI) in differentiating metastatic from non-metastatic lymph nodes in patients with lung cancer. A systematic literature search was performed to identify eligible original studies. The quality of included studies was assessed using 'quality assessment of diagnostic accuracy studies' (QUADAS-2). Meta-analysis was performed to pool sensitivity and specificity, to calculate the positive likelihood ratio (PLR), the negative likelihood ratio (NLR) and the diagnostic odds ratio (DOR), and to construct the summary receiver operating characteristic (SROC) curve. The homogeneity, threshold effect and publication bias were also investigated. Meta-regression analysis was performed to identify the sources of heterogeneity. A total of 10 studies with 11 datasets met the inclusion criteria, which comprised 796 patients with a total of 2,433 lymph nodes. The pooled diagnostic sensitivity was 0.78 [95% confidence interval (CI): 0.74-0.81] and the pooled diagnostic specificity was 0.88 (95% CI: 0.86-0.89). The PLR, NLR, and DOR were 7.11 (95% CI: 4.39-11.52), 0.24 (95% CI: 0.18-0.33), and 31.14 (95% CI: 17.32-55.98), respectively. The area under the SROC curve was 0.90. No publication bias was found (bias=-0.15, P=0.887). Notable heterogeneity was, however, observed, and patient selection, type of lung cancer, number of enrolled lymph nodes, reference standard, B-value and the type of scanner were the

sources of heterogeneity (P<0.05). No significant threshold effect was identified (P=0.537). In conclusion, DWI has been revealed to be a valuable magnetic resonance imaging (MRI) modality, with good diagnostic performance for distinguishing metastatic from non-metastatic lymph nodes in patients with lung cancer. Therefore, DWI may be a useful supplement to conventional MRI techniques.

Introduction

Globally, lung cancer has been revealed to be the most common cancer, the leading cause of cancer mortality in males, and the second leading cause of cancer mortality in females during the year 2008 (1). Although substantial progress in surgical, chemotherapeutic and radiotherapeutic approaches has been made, the long-term survival rate in patients with lung cancer still remains low (2). A key determinant of the appropriate therapy provided to patients with lung cancer depends on the accurate detection of lymph node metastasis. Therefore, the correct assessment of the lymph node status is crucial for the prognosis of patients with lung cancer.

Non-invasive diagnostic imaging modalities, i.e., computed tomography (CT) and magnetic resonance imaging (MRI), are commonly used for the identification of metastatic lymph nodes in patients with lung cancer. However, conventional CT and MRI rely on morphological criteria to differentiate malignant from benign lymph nodes. Lymph nodes may be enlarged due to other factors, including inflammation, infection, auto-immune disease, and so forth, and even small lymph nodes may be infiltrated by cancer cells (3,4). Therefore, it remains a challenge for radiologists and clinicians, on the basis of morphological characteristics, to distinguish metastatic from non-metastatic lymph nodes. Positron emission tomography (PET)/CT is a generally accepted non-invasive approach for the staging of pulmonary cancer (5,6). However, the use of PET/CT has been known to result in false-negative results for well-differentiated lung adenocarcinoma, and false-positive results for inflammatory lesions (7,8). Furthermore, PET/CT is not widely available on account of its great expense.

Recently, diffusion-weighted imaging (DWI) has successfully been applied to thoracic imaging, and has demonstrated

Correspondence to: Dr Mao-Hua Wang, Department of Anesthesiology, The Affiliated Hospital of Southwest Medical University, 25 Taiping Street, Luzhou, Sichuan 646000, P.R. China
E-mail: wmhsunny@126.com

Dr Guang-Xiang Chen, Department of Radiology, The Affiliated Hospital of Southwest Medical University, 25 Taiping Street, Luzhou, Sichuan 646000, P.R. China
E-mail: cgx23ly2002@163.com

Key words: diffusion-weighted imaging, magnetic resonance imaging, lung cancer, lymph node, meta-analysis

great potential in the detection of lung cancer and metastatic lymph nodes (5,9-15). However, previously published studies have reported the diagnostic performance of DWI for the detection of metastatic lymph nodes with widely varying sensitivities and specificities (60-91 and 70-100%, respectively) (11,14-22). Thus, there has been a burgeoning interest in meta-analysis to identify consistent results for the diagnostic performance of DWI in detecting metastatic lymph nodes of lung cancer.

Wu *et al* (23) published a meta-analysis that compared the diagnostic capability of DWI with that of ^{18}F -fluorodeoxyglucose PET/CT in the lymph node staging evaluation of non-small cell lung cancer (NSCLC); however, only three DWI studies were included in that study. Another meta-analysis of differential diagnosis between metastatic and non-metastatic lymph nodes using DWI featured a great variety of primary tumors, although this meta-analysis only included four DWI studies on lung cancer (24). The different biological or histological characteristics of each primary tumor may lead to decreased reliability of the pooled results. The aim of the present study was to perform a meta-analysis for more comprehensive and precise assessment of the diagnostic value of DWI, in order to discriminate between metastatic and non-metastatic lymph nodes in patients with lung cancer.

Materials and methods

Literature search strategy. Systematic literature searches of the PubMed, Embase, Web of Science, Cochrane Library, China Biomedicine, China National Knowledge Infrastructure and Wanfang databases were performed for English and Chinese language studies published between January 1994 and June 2015. A search algorithm based on a combination of the following terms was used: 'Diffusion magnetic resonance imaging' OR 'diffusion MR imaging' OR 'diffusion weighted imaging [MeSH]' OR 'DWI' AND 'lung neoplasms [MeSH]' OR 'lung cancer' OR 'lung lesions' AND 'lymphatic metastasis' OR 'lymph node metastasis'. The reference lists of identified articles were also manually searched to obtain additional papers.

Inclusion and exclusion criteria. All DWI studies for detecting lymph node metastasis in patients with lung cancer yielded by our document retrieval were assessed for potential eligibility. The studies that fulfilled the following inclusion criteria were included: i) The studies were published in the English or Chinese language in a peer-reviewed journal; ii) the studies must have evaluated the diagnostic performance of DWI for identifying metastatic and non-metastatic lymph nodes in patients with lung cancer; iii) the studies had to present on lymph node-based, not on patient-based, statistical results; iv) sufficient information regarding true-positive (TP), false-positive (FP), true-negative (TN) and false-negative (FN) values could be identified or calculated from data in the original articles; v) pathological examination and/or follow-up imaging findings should have been set as the reference standards; and vi) where data or subsets of data were reported in more than one article, the article with the larger sample size, or the latest article, was adopted.

Studies were excluded if they met at least one of the following deficiencies: i) The studies were conference abstracts, letters, comments, case reports or review articles; ii) the studies were not associated with lung cancer; iii) combinations of multiple magnetic resonance sequences, i.e., contrast enhanced imaging and DWI, were applied in the studies, and the diagnostic performance of DWI could not be assessed alone; iv) sufficient raw data could not be obtained from the articles to construct the four-fold (2x2) tables; or v) sample sizes of <20 patients or 30 lymph nodes were presented in the studies.

Data extraction and quality assessment. Two authors (G.-X.C and M.-H.W) independently searched the literature, screened the retrieved articles, and extracted relevant data, including the study characteristics and results using Microsoft Excel sheets. For each included study, the following characteristics were extracted: The first author's name, year of publication, age and gender of patients, number of patients and lymph nodes, type of lung cancer, study design, patient selection, blind study, B-value, diagnostic threshold of apparent diffusion coefficient (ADC), reference standard, type of scanner, MRI field strength and the relevant data (i.e., TP, FP, TN and FN). The quality of the included studies was also assessed independently by the same two authors according to the 'quality assessment of diagnostic accuracy studies' (QUADAS-2) (25). If agreement was not reached between the two authors, a third author (T.Z.) evaluated all discrepant items, and the major viewpoint was adopted for statistical analysis.

Statistical analysis. Heterogeneity was assessed using the Q statistic of the Chi-square value test and inconsistency index (I^2), and $P < 0.1$ or $I^2 > 50\%$ suggested the presence of heterogeneity (26). If heterogeneity was detected among the studies, the test's performance was summarized using a random-effects model (the DerSimonian and Laird method); otherwise, a fixed-effects model (the Mantel-Haenszel method) was used. The predominant outcome indexes were pooled sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), diagnostic odds ratio (DOR) and their corresponding 95% confidence intervals (CIs). The summary receiver operating characteristic (SROC) curve and area under the SROC curve (AUC) were used to demonstrate the diagnostic performance of DWI in the detection of metastatic lymph nodes of lung cancer.

Exploring the sources of heterogeneity is essential to seek the latent factors that affect the pooled accuracy from various studies. In diagnostic accuracy studies, the threshold effect has been regarded as one of the predominant causes of heterogeneity (27). The threshold effect can be recognized visually by observing the typical pattern of a 'shoulder-arm' shape in the ROC plane. Furthermore, Spearman's correlation coefficient between the logit of sensitivity and the logit of (1-specificity) was calculated to identify the presence of a threshold effect. Strong positive correlation, with $P < 0.05$, was considered to indicate the existence of a threshold effect (28).

In addition, in order to investigate the influence of heterogeneity between individual studies on pooled diagnostic performance, meta-regression analysis was performed to identify those variables causing heterogeneity. Subgroup analysis

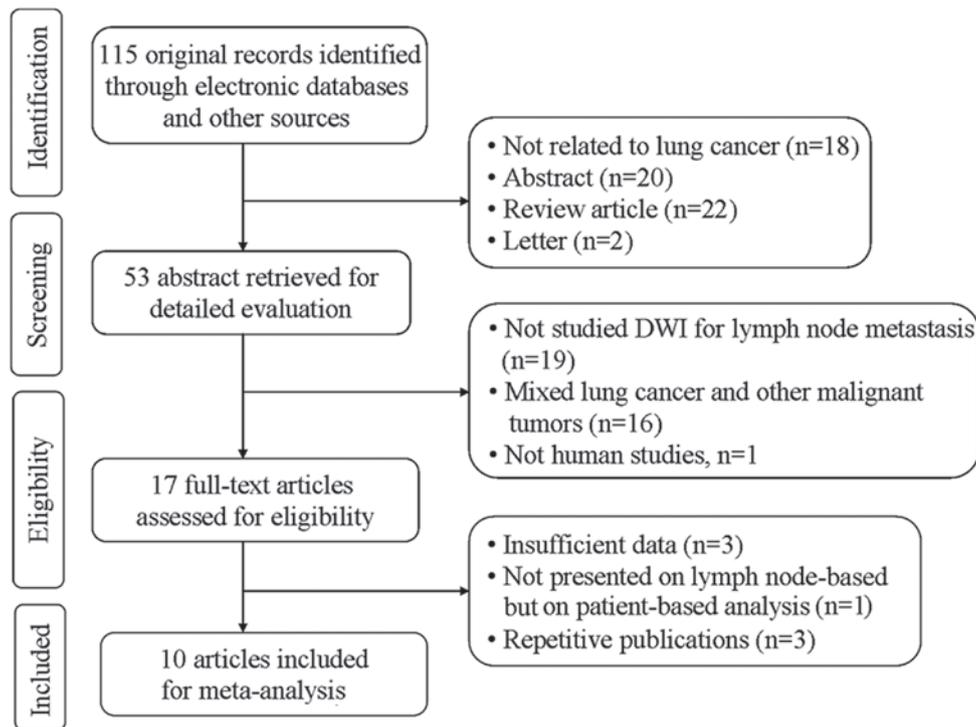


Figure 1. Flow-chart of the study selection process. DWI, diffusion-weighted imaging.

was further performed according to these identified variables. The freely available software, Meta-DiSc (version 1.4), was used to perform the heterogeneity test, assessment of threshold effect, pooling of diagnostic performance, and the meta-regression and subgroup analyses (29).

The possible existence of publication bias was visually assessed with the Deeks' funnel plot and asymmetry test, using Stata software (version 12.0; StataCorp LP, College Station, TX, USA). An inverted symmetrical funnel plot with $P > 0.05$ was considered to indicate the absence of publication bias (30).

Results

Included studies. The search strategy initially yielded 115 studies, of which 10 studies (11,14-22) with 11 datasets met the inclusion criteria, including 796 patients (472 male and 324 female; mean age, 66.3 years) with a total of 2,433 lymph nodes. In one study (15), the analyses were performed based on normal-sized lymph nodes and enlarged lymph nodes, respectively. Thus, that study was regarded as comprising two unique datasets, which were included independently in the present meta-analysis. The detailed flowchart of the study selection process is shown in Fig. 1, and the characteristics of the included studies are summarized in Table I.

Assessment of study quality and publication bias. The quality assessment of the included studies was moderate according to the QUADAS-2 tool, and the results are shown in Fig. 2. The Deeks' funnel plot asymmetry tests revealed no strong evidence for the presence of publication bias (bias=-0.15, $P=0.887$).

Homogeneity tests and threshold effect analysis. Homogeneity tests demonstrated significant evidence of heterogeneity

for sensitivity ($I^2=73.7\%$, $P<0.0001$), specificity ($I^2=95.2\%$, $P<0.0001$), PLR ($I^2=86.0\%$, $P<0.0001$), NLR ($I^2=68.6\%$, $P=0.0004$), and DOR ($I^2=66.1\%$, $P=0.0010$). The typical pattern of a 'shoulder-arm' shape in the ROC plane was not identified. Furthermore, Spearman's correlation coefficient, a further test for the threshold effect, was calculated to be 0.209 ($P=0.537$), which indicated that no threshold effect existed.

Diagnostic accuracy of DWI. Due to the notable heterogeneity among the included studies, pooled values were determined using a random-effects model analysis. The pooled diagnostic sensitivity was 0.78 (95% CI: 0.74-0.81), and the pooled diagnostic specificity was 0.88 (95% CI: 0.86-0.89). The PLR, NLR, and DOR were 7.11 (95% CI: 4.39-11.52), 0.24 (95% CI: 0.18-0.33), and 31.14 (95% CI: 17.32-55.98), respectively. The forest plots for the included studies are shown in Fig. 3. The SROC curve showed that the Q value was 0.83, while the AUC was 0.90, suggesting good diagnostic performance. The SROC curve is shown in Fig. 4.

Meta-regression and subgroup analysis. Meta-regression analysis was subsequently performed to explore other sources (except for threshold effect) of heterogeneity. The results revealed that patient selection, type of lung cancer, number of enrolled lymph nodes, reference standard, B-value and type of scanner were the sources of heterogeneity ($P<0.05$). Study design, blinding to all items of information concerning the other test results, and field strength did not statistically contribute to the heterogeneity. Table II shows the detailed results of the meta-regression analysis. The subgroup analysis was performed for the above-identified variables whose sample size was not <5 datasets. As the sample sizes of consecutive patient selection (only 4 datasets), the type including small cell

Table I. Characteristics of the included studies in the meta-analysis.

Study ID	First author	Year	Age (range/average)	Gender (M/F)	Study size ^a	Type of lung cancer	Design	Patient enrollment	Blind	Reference standard	B-value (s/mm ²)	DT of ADC (mm ² /s)	Field strength	Type of scanner	Refs.
1	Chen	2010	35-76/51	35/21	56/135	NSCLC	UN	Cons	Yes	Pathology and/or follow-up	1,000	UN	1.5 T	Siemens	(14)
2	Nakayama	2010	48-82/68	38/32	70/56	NSCLC	Retro	UN	UN	Pathology	1,000	1.54x10 ⁻³	1.5 T	Siemens	(20)
3	Nomori	2008	38-82/70	47/41	88/734	NSCLC	Pros	UN	UN	Pathology	1,000	1.60x10 ⁻³	1.5 T	Phillips	(11)
4	Usuda	2013	37-83/68	94/64	158/705	Lung cancer	UN	UN	UN	Pathology	800	1.70x10 ⁻³	1.5 T	Siemens	(18)
5	Xu (ELN)	2014	42-78/55	27/15	42/33	NSCLC	Pros	Cons	Yes	Pathology	1,000	1.98x10 ⁻³	1.5 T	Phillips	(15)
6	Xu (NLN)	2014	42-78/55	27/15	42/86	NSCLC	Pros	Cons	Yes	Pathology	1,000	2.04x10 ⁻³	1.5 T	Phillips	(15)
7	Ohno	2011	61-83/73	136/114	250/270	NSCLC	Pros	Cons	Yes	Pathology	1,000	2.50x10 ⁻³	1.5 T	Phillips	(19)
8	Bai	2013	27-72/59	16/10	26/62	NSCLC	UN	UN	UN	Pathology and/or follow-up	600	1.92x10 ⁻³	1.5 T	Phillips	(21)
9	He	2013	33-77/58	27/9	36/206	Lung cancer	UN	UN	UN	Pathology	500	3.19x10 ⁻³	1.5 T	GE	(16)
10	Zhang	2013	41-72/59	17/8	25/78	NSCLC	UN	UN	Yes	Pathology	800	2.21x10 ⁻³	3.0 T	Siemens	(17)
11	Zeng	2012	48-69/58	35/10	45/68	NSCLC	UN	UN	Yes	Pathology	600	2.32x10 ⁻³	1.5 T	GE	(22)

^aNumber of enrolled patients and lymph nodes. NSCLC, non-small cell lung cancer; ELN, enlarged lymph nodes; NLN, normal-sized lymph nodes; UN, unclear; Retro, retrospective; Pros, prospective; Cons, consecutive; GE, General Electric.

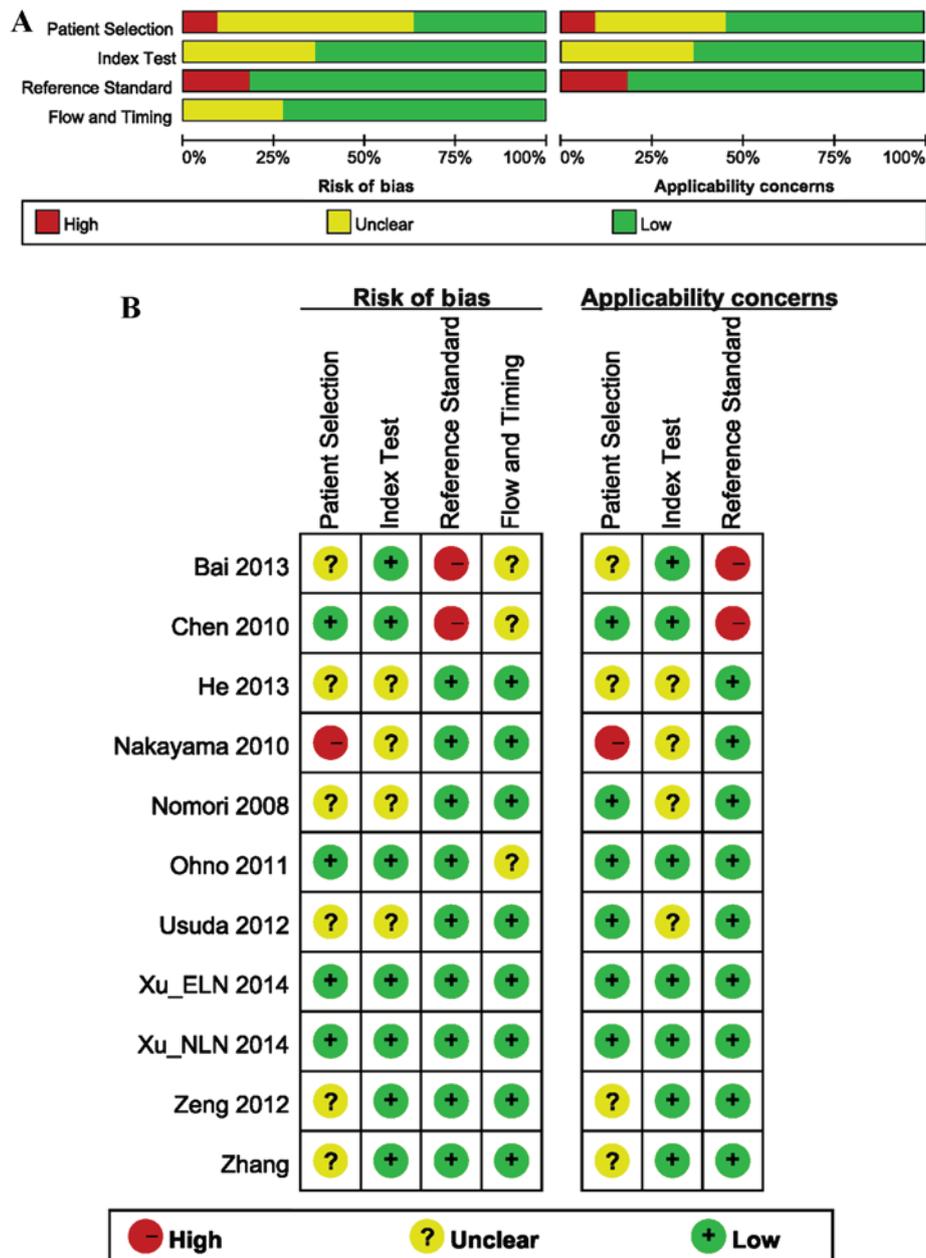


Figure 2. Methodological quality assessment of the included studies, according to QUADAS-2. (A) Overview of the entire meta-analysis. (B) Quality assessment as determined on an individual study basis. QUADAS, quality assessment of diagnostic accuracy studies.

lung cancer and reference standard of pathology and follow-up (only 2 datasets) were too small to conduct subgroup analysis, only subgroup analysis of NSCLC and the reference standard of pathology were performed among the three variables. The results of subgroup analysis of NSCLC and B-value equal to 1,000 demonstrated the highest diagnostic performance for pooled sensitivity, specificity and DOR. The results of the subgroup analysis are shown in Table III.

Discussion

The presence of lymph node metastasis in patients with lung cancer has been demonstrated to be a pivotal prognostic factor for cancer staging and the outcome in these patients (31-33). A number of studies have indicated that DWI is a promising

technique to distinguish metastatic from non-metastatic lymph nodes (11,14,15,20). In the present study, the diagnostic performance of DWI in detecting metastatic lymph nodes of lung cancer was investigated. The results revealed that, for lymph node metastasis detection, DWI had relatively low sensitivity (78%) and high specificity (88%). The SROC curve presents a global summary of test performance, and displays the trade-off between sensitivity and specificity (34-36). The AUC was revealed to be 0.90, which indicated a good diagnostic ability. The DOR derived from different combinations of sensitivity and specificity is the ratio of the odds of positivity in the diseased state relative to the odds of positivity in the non-diseased state, and may be used as a single summary measure. The value of a DOR ranges from 0 to infinity, and the greater it is, the better it is able to distinguish test

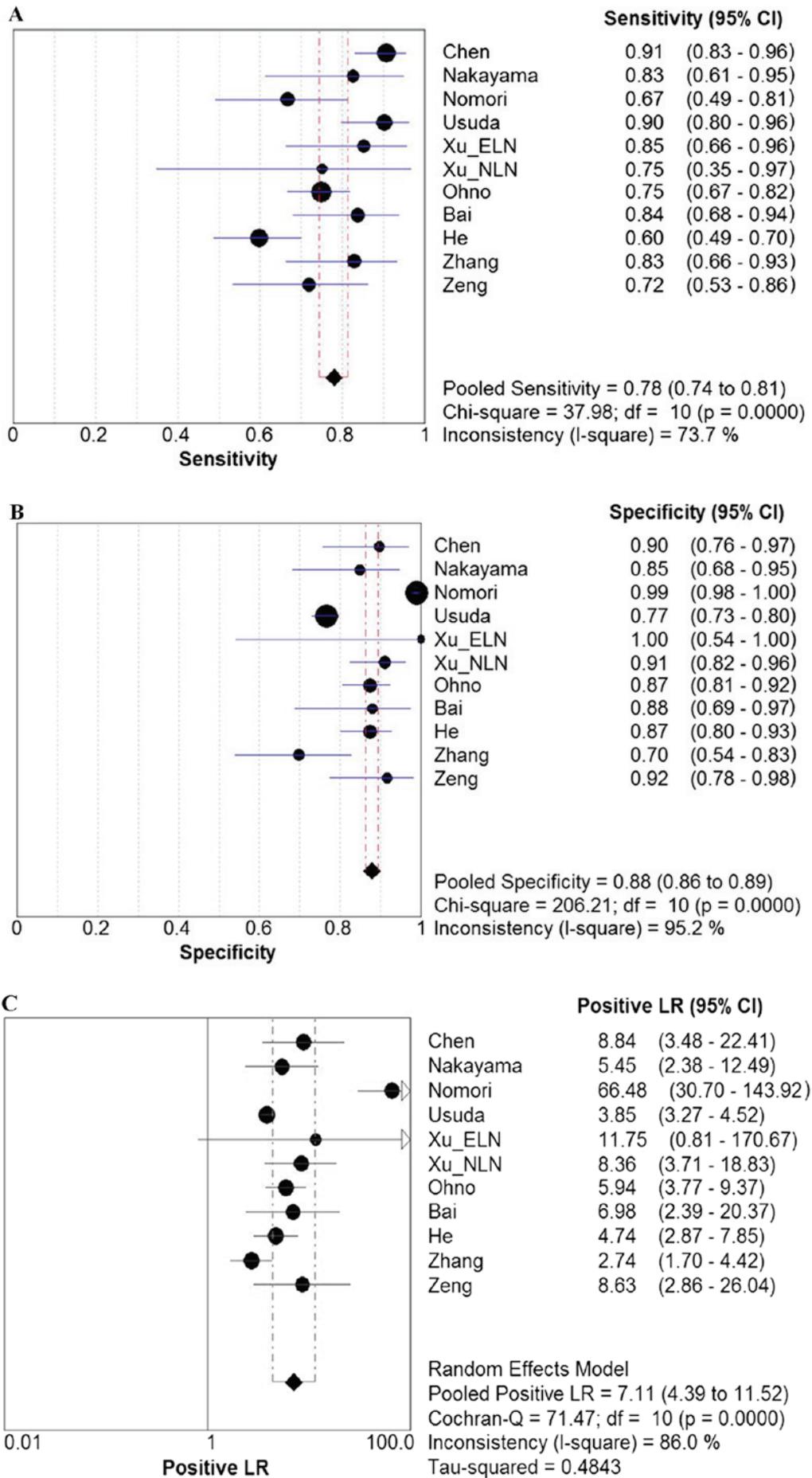


Figure 3. Forest plots. Forest plots of the (A) sensitivity, (B) specificity, (C) PLR.

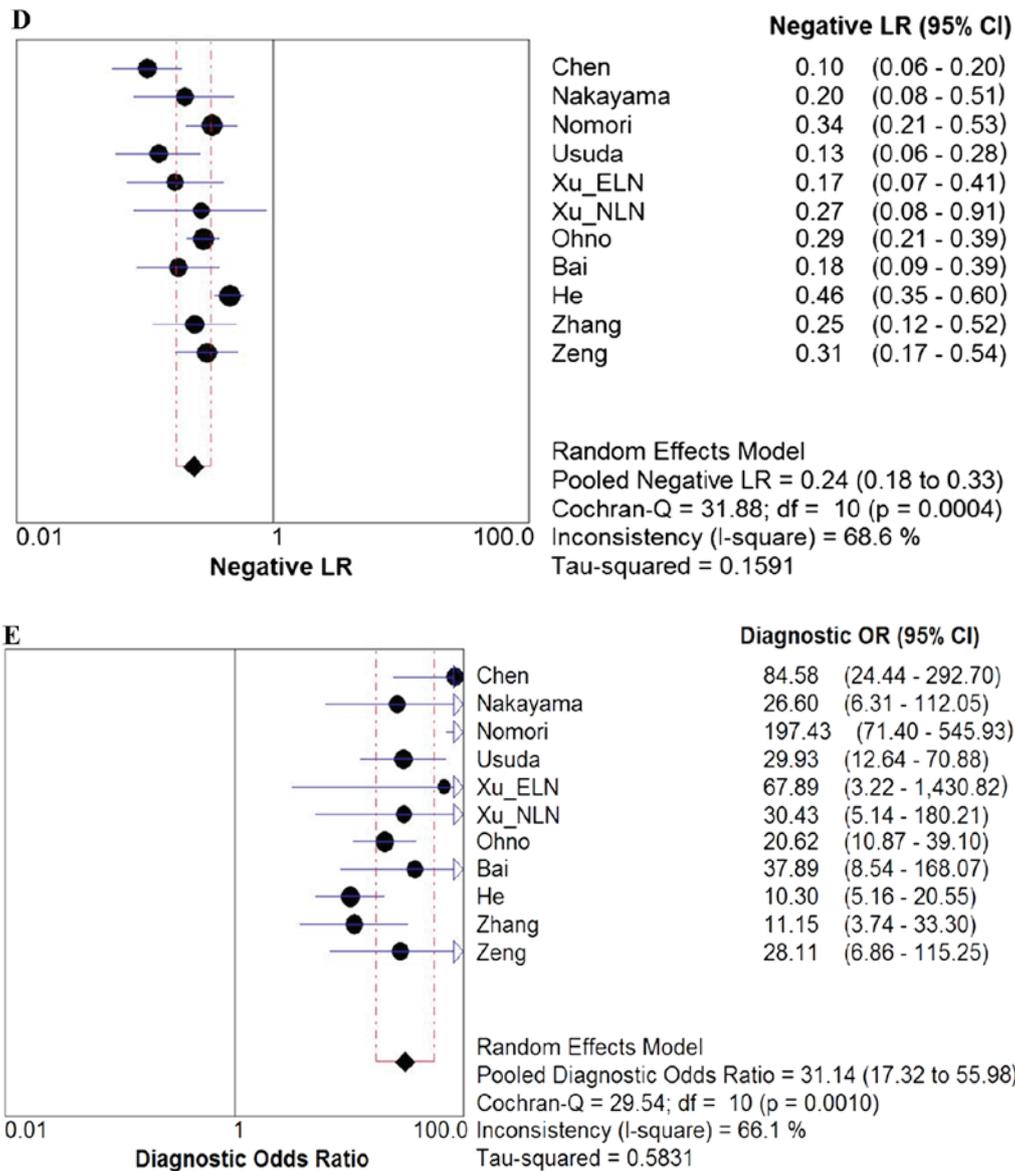


Figure 3. Continued. (D) NLR and (E) DOR with corresponding 95% CIs for DWI in detection of metastatic lymph nodes of lung cancer from all included studies are shown. QUADAS, quality assessment of diagnostic accuracy studies; PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; CIs, confidence intervals; DWI, diffusion-weighted imaging.

performance (23,37). In the present study, it was identified that the DOR for DWI in detecting metastatic lymph nodes of lung cancer was 31.14 (95% CI: 17.32-55.98), which also indicated high overall diagnostic accuracy.

In clinical practice, it is necessary to know how a diagnostic test result predicts the risk of abnormality. The likelihood ratio (LR) is generally considered to be one of the best measures of diagnostic accuracy, and is even more helpful for decision-making in clinical practice compared with measures of sensitivity, specificity or the AUC (24). An LR >1 is indicative of the fact that the test result is associated with the presence of the disease, whereas, by contrast, an LR <1 indicates that the test result is associated with the absence of disease. The more that the LR deviates from a value of 1, the stronger is the proof for either the presence or absence of disease (38). A higher PLR lends itself to a ruling in favor of a disease, whereas a lower NLR would

lend itself to ruling out the possibility of disease. In order to be moderately useful, a diagnostic test should have a higher PLR value (>5) and a lower NLR value (<0.2) (39). In the present study, the pooled PLR and NLR were 7.11 and 0.24, respectively, which demonstrated that DWI had a moderately good diagnostic performance for ruling in cases of lymph node metastasis in patients with lung cancer, but a relatively inferior diagnostic ability for ruling out non-metastatic lymph nodes. Therefore, a lung cancer patient with lymph node metastasis was 7.11 times more likely to have a positive DWI manifestation compared with a patient without lymph node metastasis. Analogously, the probability of having a negative DWI manifestation for lymph node metastasis in patients with lung cancer was 0.24 times (or approximately one-fourth) that of those without lymph node metastasis. In other words, non-metastatic lymph nodes were approximately 4 times more likely to have a negative DWI manifestation compared with

Table II. Results of meta-regression analysis.

Variable	Coefficient	Standard error	P-value	RDOR	95% CI
Design	-0.629	0.5875	0.3630	0.53	0.08-3.46
Patient selection	-1.824	0.5447	0.0154	0.16	0.04-0.61
Blind	-1.455	1.0608	0.2420	0.23	0.01-4.44
Reference standard	-1.752	0.5831	0.0239	0.17	0.04-0.72
Type of lung cancer	1.270	0.4376	0.0198	3.56	1.30-9.76
Study size ^a	0.003	0.0008	0.0138	1.00	1.00-1.00
B-value	0.007	0.0016	0.0050	1.01	1.00-1.01
DT of ADC	-4.786	2.0793	0.0828	0.01	0.00-2.68
Type of scanner	0.970	0.3630	0.0442	2.64	1.04-6.71
Field strength	-0.853	0.4270	0.0926	0.43	0.15-1.21

^aNumber of enrolled lymph nodes. RDOR, relative diagnostic odd ratio; CI, confidence interval; DT, diagnostic threshold; ADC, apparent diffusion coefficient.

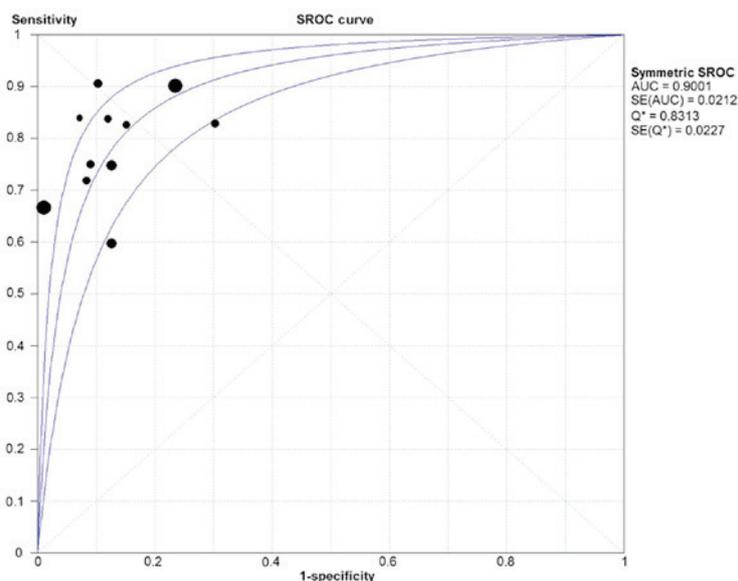


Figure 4. SROC curve for DWI in detection of metastatic lymph nodes of lung cancer from all included studies. SROC, summary receiver operating characteristic; DWI, diffusion-weighted imaging; AUC, area under the curve.

metastatic lymph nodes in patients with lung cancer. Owing to the larger number of the included studies in the present meta-analysis, these results were not entirely consistent with previous studies (23,24). However, the results in the present study do circumvent several of the limitations that were acknowledged in previous studies, and provide objective and practical suggestions for DWI in the evaluation of lymph node status in patients with lung cancer.

Significant heterogeneity between the included studies was identified in the present meta-analysis. Exploring the sources of heterogeneity, which is useful in order to understand the potential factors that influence accuracy assessments of the pooled diagnostic performance, is an important aim of meta-analysis (40,41). In the present study, the threshold effect was assessed using the ROC plane and Spearman's correlation coefficient. As indicated by the results of threshold effect analysis, there was no threshold effect, and therefore

the heterogeneity may have been caused by other factors, for example, the study characteristics. Thus, meta-regression analysis was performed to further explore the sources of heterogeneity. The results demonstrated that patient selection, the type of lung cancer, the number of enrolled lymph nodes, the reference standard, B-value and type of scanner were strongly associated with diagnostic accuracy. In the subgroup analysis based on the type of lung cancer and the B-value, the results of the subgroup analysis of NSCLC and the B-value equal to 1,000 demonstrated the highest diagnostic performance for pooled sensitivity, specificity and DOR. The results further verified that the type of lung cancer involved and the B-value are able to affect the diagnostic accuracy of DWI, and therefore, the heterogeneity in these studies was increased. In the present study, the two studies did not distinguish the type of lung cancer, i.e. they did not differentiate NSCLC from small-cell lung carcinoma. Previous studies have revealed

Table III. Results of subgroup analysis for DWI in detecting lymph node metastasis.

	No. of studies	Pooled sensitivity		Pooled specificity		DOR	
		Value (95% CI)	I ²	Value (95% CI)	I ²	Value (95% CI)	I ²
Total ^a	11	0.78 (0.74-0.81)	73.0	0.88 (0.86-0.89)	95.2	31.14 (17.32-55.98)	66.1
Type of lung cancer							
NSCLC	9	0.80 ^b (0.76-0.85)	50.5 ^c	0.95 ^b (0.93-0.96)	90.7 ^c	37.71 ^b (19.03-74.78)	61.6 ^c
Study size ^d							
≥100	5	0.77 (0.73-0.81)	88.4	0.88 (0.87-0.90)	97.9	37.66 ^b (14.17-100.07)	85.0
<100	6	0.81 ^b (0.74-0.87)	0.0 ^c	0.86 (0.81-0.90)	59.5 ^c	22.96 (12.49-42.23)	0.0 ^c
Reference standard							
Pathology	9	0.75 (0.71-0.79)	66.5 ^c	0.88 (0.86-0.89)	96.1	27.46 (14.30-52.71)	68.7
B-value							
1,000	6	0.80 ^b (0.75-0.84)	65.3 ^c	0.96 ^b (0.95-0.97)	90.6 ^c	50.73 ^b (19.91-129.26)	68.9
<1,000	5	0.75 (0.70-0.81)	81.7	0.79 (0.76-0.81)	74.3 ^c	18.04 (10.31-31.55)	32.1 ^c
Type of scanner							
Phillips	5	0.76 (0.70-0.81)	9.7 ^c	0.96 ^b (0.95-0.98)	91.4 ^c	48.84 ^b (15.90-149.98)	72.5
Siemens and GE	6	0.79 ^b (0.75-0.84)	84.7	0.79 (0.76-0.82)	72.5 ^c	22.54 (11.80-43.06)	54.8 ^c

^aDiagnostic accuracy and heterogeneity of all 11 included datasets; ^bthe subgroup of higher diagnostic performance compared with the total; ^cthe subgroup of lower heterogeneity compared with the total; ^dnumber of enrolled lymph nodes. CI, confidence interval; NSCLC, non-small cell lung cancer; DOR, diagnostic odds ratio; GE, General Electric.

that the mean ADC value of small-cell carcinoma is significantly different from that of NSCLC (42,43). In addition, the selection of the B-value is particularly important for DWI. At lower B-values, with <600 sec/mm², signals are prominently influenced by perfusion effects. Thus, their ADC values may not reflect diffusion phenomena alone. To prevent perfusion effects, it is necessary to use higher B-values. However, the higher the B-value, the greater is the chance of distortion and susceptibility artifacts occurring (44,45). The present study revealed that different B-values contributed to adding to the heterogeneity, and studies of the B-value equal to 1,000 had the best diagnostic performance. These findings are helpful for the clinical application of DWI in detecting lymph node metastasis of lung cancer.

Although the present meta-analysis has produced a more robust assessment of the true effect-size, with less random error compared with individual studies, it should be recognized that our study had several limitations. First, although no publication bias was found in our meta-analysis, publication bias may potentially still exist. The present meta-analysis was based exclusively on published studies, and a grey literature search analysis was not performed, which may have led to an overestimation of the true effect. Additionally, the present study only included previously published studies that were written in English or Chinese, which may also have introduced unavoidable inclusion bias. However, this bias would have been likely to have been small, since the majority of studies of high quality were published in English. Secondly, there was a notable heterogeneity among the included studies. Although meta-regression analysis was adopted to explore the sources of heterogeneity, the analysis of heterogeneity may still have been insufficient. The optimal acquisition protocol of DWI has

not been defined, and differences attributable to the different manufacturer of the MRI scanner, magnetic field strengths and parameters used in DWI series are capable of affecting the quality of DWI. In the present study, meta-regression analysis and subgroup analysis were utilized to identify variables that may be responsible for the heterogeneity. However, it is impossible to perform meta-regression and subgroup analyses for all variables, due to the excessive number of factors that are involved in the processes of DWI. Therefore, it is necessary to further develop a standard acquisition protocol of DWI as a routine clinical application. Thirdly, among the included studies, there were numerous non-prospective studies, and certain of the studies were designed without a declaration of DWI reviewer-blinding to all items of information concerning the results of other examinations. Furthermore, the reference standards used in two included studies were pathology and/or clinical follow-up, and these studies did not report the exact number of metastatic lymph nodes that had been diagnosed by pathology or follow-up. A reference standard based on clinical follow-up can lead to inaccuracies in the sensitivities or specificities provided by that study, and therefore the credibility of the results is reduced.

In conclusion, our meta-analysis has demonstrated that DWI is a valuable MRI modality, with good diagnostic performance for distinguishing metastatic from non-metastatic lymph nodes in patients with lung cancer. Therefore, DWI has been shown to be a useful supplement to conventional MRI techniques. In the future, larger-scale prospective studies with respect to DWI for the diagnosis of lymph node metastasis will be required to evaluate and confirm its clinical value. Furthermore, optimization of the DWI acquisition protocol, standard image processing and analysis are crucial for the

routine clinical application of DWI in detecting lymph node metastasis in patients with lung cancer.

Acknowledgements

The present study was supported by the National Natural Science Foundation of China (grant no. 81541090) and the Joined Foundation of Luzhou Municipal Government and Southwest Medical University [2015LZCYD-S04 (9/15)].

References

- Jemal A, Bray F, Center MM, Ferlay J, Ward E and Forman D: Global cancer statistics. *CA Cancer J Clin* 61: 69-90, 2011.
- Jemal A, Siegel R, Xu J and Ward E: Cancer statistics, 2010. *CA Cancer J Clin* 60: 277-300, 2010.
- Kligerman S and Digumarthy S: Staging of non-small cell lung cancer using integrated PET/CT. *AJR Am J Roentgenol* 193: 1203-1211, 2009.
- De Wever W: Role of integrated PET/CT in the staging of non-small cell lung cancer. *JBR-BTR* 92: 124-126, 2009.
- Pauls S, Schmidt SA, Juchems MS, Klass O, Luster M, Reske SN, Brambs HJ and Feuerlein S: Diffusion-weighted MR imaging in comparison to integrated [¹⁸F]-FDG PET/CT for N-staging in patients with lung cancer. *Eur J Radiol* 81: 178-182, 2012.
- Kim YN, Yi CA, Lee KS, Kwon OJ, Lee HY, Kim BT, Choi JY, Kim SW, Chung MP, Han J, *et al*: A proposal for combined MRI and PET/CT interpretation criteria for preoperative nodal staging in non-small-cell lung cancer. *Eur Radiol* 22: 1537-1546, 2012.
- Cheran SK, Nielsen ND and Patz EF Jr: False-negative findings for primary lung tumors on FDG positron emission tomography: Staging and prognostic implications. *AJR Am J Roentgenol* 182: 1129-1132, 2004.
- Shim SS, Lee KS, Kim BT, Choi JY, Chung MJ and Lee EJ: Focal parenchymal lung lesions showing a potential of false-positive and false-negative interpretations on integrated PET/CT. *AJR Am J Roentgenol* 186: 639-648, 2006.
- Usuda K, Zhao XT, Sagawa M, Aikawa H, Ueno M, Tanaka M, Machida Y, Matoba M, Ueda Y and Sakuma T: Diffusion-weighted imaging (DWI) signal intensity and distribution represent the amount of cancer cells and their distribution in primary lung cancer. *Clin Imaging* 37: 265-272, 2013.
- Ohba Y, Nomori H, Mori T, Ikeda K, Shibata H, Kobayashi H, Shiraishi S and Katahira K: Is diffusion-weighted magnetic resonance imaging superior to positron emission tomography with fludeoxyglucose F 18 in imaging non-small cell lung cancer? *J Thorac Cardiovasc Surg* 138: 439-445, 2009.
- Nomori H, Mori T, Ikeda K, Kawanaka K, Shiraishi S, Katahira K and Yamashita Y: Diffusion-weighted magnetic resonance imaging can be used in place of positron emission tomography for N staging of non-small cell lung cancer with fewer false-positive results. *J Thorac Cardiovasc Surg* 135: 816-822, 2008.
- Nomori H, Cong Y, Abe M, Sugimura H and Kato Y: Diffusion-weighted magnetic resonance imaging in preoperative assessment of non-small cell lung cancer. *J Thorac Cardiovasc Surg* 149: 991-996, 2015.
- Koyama H, Ohno Y, Nishio M, Takenaka D, Yoshikawa T, Matsumoto S, Seki S, Maniwa Y, Ito T, Nishimura Y and Sugimura K: Diffusion-weighted imaging vs STIR turbo SE imaging: Capability for quantitative differentiation of small-cell lung cancer from non-small-cell lung cancer. *Br J Radiol* 87: 20130307, 2014.
- Chen W, Jian W, Li HT, Li C, Zhang YK, Xie B, Zhou DQ, Dai YM, Lin Y, Lu M, *et al*: Whole-body diffusion-weighted imaging vs. FDG-PET for the detection of non-small-cell lung cancer. How do they measure up? *Magn Reson Imaging* 28: 613-620, 2010.
- Xu L, Tian J, Liu Y and Li C: Accuracy of diffusion-weighted (DW) MRI with background signal suppression (MR-DWIBS) in diagnosis of mediastinal lymph node metastasis of non-small-cell lung cancer (NSCLC). *J Magn Reson Imaging* 40: 200-205, 2014.
- He W, Xu JP, Zhou XH, *et al*: Comparison of CT and DWI in preoperative evaluation of chest lymph node status in lung cancer. *Journal of Clinical Radiology* 32: 802-806, 2013 (In Chinese).
- Zhang X, Xing W and Chen J: Application of DWI in differential diagnosis of lymph nodes in patients with lung cancer. *Chin Comput Med Imag* 19: 213-216, 2013 (In Chinese).
- Usuda K, Sagawa M, Motono N, Ueno M, Tanaka M, Machida Y, Matoba M, Kuginuki Y, Taniguchi M, Ueda Y and Sakuma T: Advantages of diffusion-weighted imaging over positron emission tomography-computed tomography in assessment of hilar and mediastinal lymph node in lung cancer. *Ann Surg Oncol* 20: 1676-1683, 2013.
- Ohno Y, Koyama H, Yoshikawa T, Nishio M, Aoyama N, Onishi Y, Takenaka D, Matsumoto S, Maniwa Y and Nishio W: N stage disease in patients with non-small cell lung cancer: Efficacy of quantitative and qualitative assessment with STIR turbo spin-echo imaging, diffusion-weighted MR imaging, and fluorodeoxyglucose PET/CT. *Radiology* 261: 605-615, 2011.
- Nakayama J, Miyasaka K, Omatsu T, Onodera Y, Terae S, Matsuno Y, Cho Y, Hida Y, Kaga K and Shirato H: Metastases in mediastinal and hilar lymph nodes in patients with non-small cell lung cancer: Quantitative assessment with diffusion-weighted magnetic resonance imaging and apparent diffusion coefficient. *J Comput Assist Tomogr* 34: 1-8, 2010.
- Bai CG, Zhang XM and Qiao W: Application of apparent diffusion coefficient in evaluating lymphatic metastasis of non-small cell lung cancer. *Jiangsu Med J* 39: 2977-2979, 2013 (In Chinese).
- Zeng Z, Liao Q, Cai J and Liu A: Diffusion-weighted imaging and apparent diffusion coefficient values in the differential diagnosis of hilar and mediastinal lymph nodes of non-small cell lung cancer. *Chinese Journal of Clinical Oncology* 39: 706-710, 2012 (In Chinese).
- Wu LM, Xu JR, Gu HY, Hua J, Chen J, Zhang W, Haacke EM and Hu J: Preoperative mediastinal and hilar nodal staging with diffusion-weighted magnetic resonance imaging and fluorodeoxyglucose positron emission tomography/computed tomography in patients with non-small-cell lung cancer: Which is better? *J Surg Res* 178: 304-314, 2012.
- Zhou M, Lu B, Lv G, Tang Q, Zhu J, Li J and Shi K: Differential diagnosis between metastatic and non-metastatic lymph nodes using DW-MRI: A meta-analysis of diagnostic accuracy studies. *J Cancer Res Clin Oncol* 141: 1119-1130, 2015.
- Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA and Bossuyt PM; QUADAS-2 Group: QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155: 529-536, 2011.
- Higgins JP and Thompson SG: Quantifying heterogeneity in a meta-analysis. *Stat Med* 21: 1539-1558, 2002.
- Chen LH, Zhang J, Bao J, Zhang L, Hu X, Xia Y and Wang J: Meta-analysis of diffusion-weighted MRI in the differential diagnosis of lung lesions. *J Magn Reson Imaging* 37: 1351-1358, 2013.
- Arends LR, Hamza TH, van Houwelingen JC, Heijnenbroek-Kal MH, Hunink MG and Stijnen T: Bivariate random effects meta-analysis of ROC curves. *Med Decis Making* 28: 621-638, 2008.
- Zamora J, Abraira V, Muriel A, Khan K and Coomarasamy A: Meta-DiSc: A software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 6: 31, 2006.
- Song F, Khan KS, Dinnies J and Sutton AJ: Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int J Epidemiol* 31: 88-95, 2002.
- Herneth AM, Mayerhoefer M, Scherthaner R, Ba-Ssalamah A, Czerny Ch and Fruehwald-Pallamar J: Diffusion weighted imaging: Lymph nodes. *Eur J Radiol* 76: 398-406, 2010.
- Harders SW, Madsen HH, Hjorthaug K, Arveschoug AK, Rasmussen TR, Meldgaard P, Hoejbjerg JA, Pilegaard HK, Hager H, Rehling M and Rasmussen F: Mediastinal staging in non-small-cell lung carcinoma: Computed tomography versus F-18-fluorodeoxyglucose positron-emission tomography and computed tomography. *Cancer Imaging* 14: 23, 2014.
- Al-Sarraf N, Gately K, Lucey J, Wilson L, McGovern E and Young V: Lymph node staging by means of positron emission tomography is less accurate in non-small cell lung cancer patients with enlarged lymph nodes: Analysis of 1,145 lymph nodes. *Lung Cancer* 60: 62-68, 2008.
- Wu LM, Xu JR, Hua J, Gu HY, Chen J, Haacke EM and Hu J: Can diffusion-weighted imaging be used as a reliable sequence in the detection of malignant pulmonary nodules and masses? *Magn Reson Imaging* 31: 235-246, 2013.
- Harbord RM, Deeks JJ, Egger M, Whiting P and Sterne JA: A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 8: 239-251, 2007.

36. Rutter CM and Gatsonis CA: A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 20: 2865-2884, 2001.
37. Glas AS, Lijmer JG, Prins MH, Bossel GJ and Bossuyt PM: The diagnostic odds ratio: A single indicator of test performance. *J Clin Epidemiol* 56: 1129-1135, 2003.
38. Deeks JJ and Altman DG: Diagnostic tests 4: Likelihood ratios. *BMJ* 329: 168-169, 2004.
39. Cronin P, Dwamena BA, Kelly AM, Bernstein SJ and Carlos RC: Solitary pulmonary nodules and masses: A meta-analysis of the diagnostic utility of alternative imaging tests. *Eur Radiol* 18: 1840-1856, 2008.
40. Naaktgeboren CA, van Enst WA, Ochodo EA, de Groot JA, Hooft L, Leeftang MM, Bossuyt PM, Moons KG and Reitsma JB: Systematic overview finds variation in approaches to investigating and reporting on sources of heterogeneity in systematic reviews of diagnostic studies. *J Clin Epidemiol* 67: 1200-1209, 2014.
41. Lijmer JG, Bossuyt PM and Heisterkamp SH: Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 21: 1525-1537, 2002.
42. Matoba M, Tonami H, Kondou T, Yokota H, Higashi K, Toga H and Sakuma T: Lung carcinoma: Diffusion-weighted MR imaging-preliminary evaluation with apparent diffusion coefficient. *Radiology* 243: 570-577, 2007.
43. Liu HD, Liu Y, Yu TL and Ye N: Usefulness of diffusion-weighted MR imaging in the evaluation of pulmonary lesions. *Eur Radiol* 20: 807-815, 2010.
44. Jezzard P, Barnett AS and Pierpaoli C: Characterization of and correction for eddy current artifacts in echo planar diffusion imaging. *Magn Reson Med* 39: 801-812, 1998.
45. Bastin ME: Correction of eddy current-induced artefacts in diffusion tensor imaging using iterative cross-correlation. *Magn Reson Imaging* 17: 1011-1024, 1999.