Thoracic Cancer ISSN 1759-7706

MINI REVIEW

Magnetic resonance imaging for N staging in non-small cell lung cancer: A systematic review and meta-analysis

Yuanyuan Zhang[1], Qin Qin[2], Baosheng Li[1,2], Juan Wang[1] & Kun Zhang[3]

1 Department of Radiation Oncology, Shandong Cancer Hospital, School of Medicine and Life Sciences, University of Jinan-Shandong Academy of Medical Sciences, Jinan, Shandong Province, China
2 Department of Radiation Oncology, Shandong Cancer Hospital, Shandong University, Jinan, Shandong Province, China
3 Department of Radiology, Hohhot First Hospital, Hohhot, Inner Mongolia, China

Keywords
Lymph node staging; magnetic resonance imaging; meta-analysis; non-small cell lung cancer; systematic review.

Correspondence
Baosheng Li, Department of Radiation Oncology, Shandong Cancer Hospital, Jiyan Road 440, Jinan 250117, Shandong Province, China.
Tel: +86 531 67626161
Fax: +86 531 87984079
Email: Baoshli@yahoo.com

Received: 16 September 2014;
Accepted: 31 October 2014.

doi: 10.1111/1759-7714.12203

Thoracic Cancer **6** (2015) 123–132

Abstract

Background: Lymph node staging in non-small cell lung cancer (NSCLC) is essential for deciding appropriate treatment. This study systematically reviews the literature regarding the diagnostic performance of magnetic resonance imaging (MRI) in lymph node staging of patients with NSCLC, and determines its pooled sensitivity and specificity.

Methods: PubMed and Embase databases and the Cochrane library were used to search for relevant studies. Two reviewers independently identified the methodological quality of each study. A meta-analysis of the reported sensitivity and specificity of each study was performed.

Results: Nine studies were included. These studies had moderate to good methodological quality. Pooled sensitivity, specificity, positive likelihood ratio (LR+), negative likelihood ratio (LR–) and diagnosis odds ratio (DOR) for per-patient based analyses (7 studies) were 74%, 90%, 7.5, 0.26, and 36.7, respectively, and those for per-lymph node based analyses (5 studies) were 77%, 98%, 42.24, 0.21, and 212.35, respectively. For meta-analyses of quantitative short time inversion recovery imaging (STIR) and diffusion-weighted imaging (DWI), pooled sensitivity and specificity were 84% and 91%, and 69% and 93%, respectively. Pooled LR+ and pooled LR– were 8.44 and 0.18, and 8.36 and 0.36, respectively. The DOR was 56.29 and 27.2 respectively.

Conclusion: MRI showed high specificity in the lymph node staging of NSCLC. Quantitative STIR has greater DOR than quantitative DWI. Large, direct, and prospective studies are needed to compare the diagnostic power of STIR *versus* DWI; consistent diagnostic criteria should be established.

Introduction

Lung cancer is the leading cause of cancer death in both men and women worldwide.[1] Non-small cell lung cancer (NSCLC) accounts for 85% of all lung cancer cases.[2] Optimal treatment of patients with NSCLC depends on accurate staging, which is based on the tumor node metastasis (TNM) classification defined by the American Joint Committee on Cancer.[3]

N staging evaluates the lymph node status of hilum, mediastinum, and supraclavicular region. Node involvement within the ipsilateral peribronchial region or hilum indicates N1 disease, ipsilateral, mediastinal or subcarinal lymphadenopathy constitutes N2 disease, and metastatic contralateral, mediastinal, scalene or supraclavicular nodes represent N3 disease.[3] Accurate lymph node assessment is essential for both pre-operative staging, which may change the treatment decision, and target field delineation for patients treated with radiotherapy.[4,5]

A variety of diagnostic techniques can be performed for N staging in NSCLC. Invasive methods, which provide pathologic diagnosis of mediastinal lymph nodes, include mediastinoscopy, mediastinotomy, endobronchial ultrasound (EBUS), esophageal ultrasound (EUS) and computed tomography (CT)-guided biopsy, which are the gold standards for staging.[6,7] Noninvasive techniques include contrast-enhanced helical chest CT scans, positron emission tomography (PET), PET/CT, and magnetic resonance imaging

(MRI).[8–11] Traditionally, CT has been the primary imaging technique for the diagnosis and staging of patients with lung cancer, but it was limited by relatively low sensitivity and specificity.[10,11] Nowadays, PET/CT is widely used for the staging of lung cancer. Many articles have documented the superiority of PET/CT over PET or CT alone.[11–14]

Magnetic resonance imaging for N staging in NSCLC has been investigated since the 1980s.[15–46] With rapid technical improvements, such as multichannel MRI, systems with powerful gradients, and the development of innovative pulse sequence techniques implementing echoplanar imaging as well as parallel imaging, more and more research has investigated the diagnostic power of MRI for N staging in NSCLC. Many studies have shown high sensitivity and/or specificity of MRI for assessing lymph node metastasis.[27–46] In this article, we performed a systematic review and meta-analysis to evaluate the diagnostic power of MRI for N-staging of NSCLC.

Methods

Literature search

Two reviewers independently conducted the systematic literature search using Pubmed, Embase, and the Cochrane library. The following keywords were used: ("lung neoplasms" or "lung cancer" or "non-small cell lung cancer") AND ("MRI scan" or "magnetic resonance imaging") AND ("neoplasm staging" or "N staging" or "lymph node") AND ("sensitivity" or "specificity" or "accuracy").

The last search was carried out on 20 December 2013, with no restriction on publication date, but confined to English language articles. An additional manual search was also performed using references from retrieved articles.

Selections of studies

Studies were included according to the following criteria: (i) published in English with full-text available; (ii) evaluating NSCLC patients without extrathoracic metastasis or with potential for surgical cure; (iii) using pathological diagnosis as the standard reference; (iv) focused on the N-staging power of MRI; (v) MRI was performed with 1.5T or 3.0T magnet; (vi) sufficient data could be extracted to form a 2×2 table, either per-patient based or per-lymph node based, or both; and (vii) more than 40 patients were included in the study.

Studies were excluded if they included small cell lung cancer or if only hilar or distant lymph nodes were evaluated. Disagreements between the two reviewers were resolved by consensus.

Data extraction and methodological quality assessment

Two reviewers independently extracted relevant data about the design and results of each study. For each study, we constructed a 2×2 contingency table consisting of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) results according to MRI and reference standard. We calculated sensitivity as TP/(TP+FN), specificity as TN/(FP+TN), and diagnosis odds ratio (DOR) as $(TP \times TN)/(FP \times FN)$. For studies comparing different MRI techniques, such as quantitative and qualitative methods, diffusion-weighted imaging (DWI), and short time inversion recovery imaging (STIR), we extracted data for each technique. The methodological quality of the selected studies was assessed using quality items derived from the Quality Assessment of Diagnostic Accuracy Studies tool,[47] recommended by Cochrane Collaboration.[48]

Statistical analysis

Statistical analyses were performed using Meta-Disc version 1.4 (Unit of Clinical Biostatics, the Ramóny Cajal Hospital, Madrid, Spain) and MIDAS module for STATA version 12 (StataCorp, College Station, TX, USA). Likelihood ratio I^2 index and χ^2 tests were used to assess the heterogeneity of included studies. The I^2 index is a measure of the percentage of total variation across studies as a result of heterogeneity. If it is greater than 50%, it suggests that there is more heterogeneity between studies than would be expected to occur by chance alone. For the likelihood ratio χ^2 test, all P values < 0.05 were considered to indicate that there was heterogeneity present between studies. If heterogeneity existed, a random effects model was used for the primary meta-analysis to obtain a summary estimate for the test sensitivity with 95% confidence intervals (CIs).

Results

Literature search

A total of 164 relevant studies were identified. Nineteen studies were excluded for duplication, and 113 studies were excluded after reviewing the title and abstract. The remaining 32 studies were searched for full text, and 23 studies were further excluded for not meeting the inclusion criteria.[15–37] Figure 1 shows a detailed flowchart.

Study description and study quality

Nine studies were identified for meta-analysis,[38–46] seven studies[38,39,41,43–46] included 800 patients for per-patient data, and five studies[38–40,42,46] included 3316 lymph nodes in 489

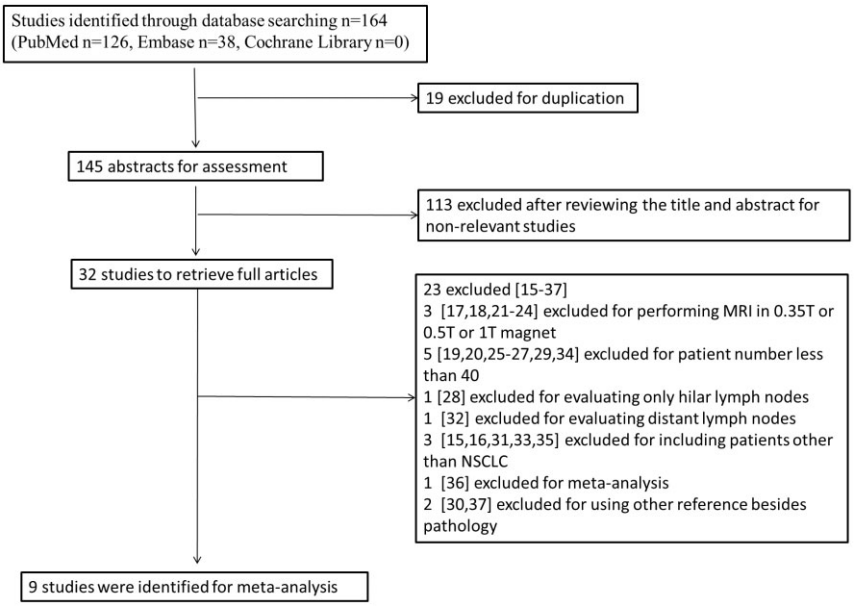

Figure 1 Flow chart of selection of studies. MRI, magnetic resonance imaging; NSCLC, non-small cell lung cancer.

patients for per-lymph node data. Detailed information on study characteristics is presented in Table 1. Considering the complexity of the MRI technique, Table 2 summarizes the pulse sequences and diagnostic criteria conducted in each study.

Table 3 exhibits the methodological quality assessment using 11 items for each of the eight included studies. All of the studies used pathological diagnosis as a reference; lymph node resection was performed either by mediastinoscopy or thoracotomy. However, there was great discrepancy among the diagnostic criteria of MRI used, which increased the heterogeneity among studies. Only three studies reported a delay between MRI examination and pathological confirmation. No study described blind measurements of reference tests without knowledge of MRI, and most studies did not provide clinical information when interpreting MRI.

Magnetic resonance imaging diagnostic accuracy

Extracted data from each study and the pooled results are presented in Figures 2 and 3. For studies comparing different MRI techniques, the data of the most accurate method was used for pooled analysis, as it represented the potential diagnostic power of MRI.

Among the studies with per-patient based data (Fig 2), the sensitivity of MRI ranged between 52% and 93%, and the specificity ranged between 82% and 100%. The pooled per-patient based sensitivity for MRI was 74% (95% CI, 69–79%), and the specificity was 90% (95% CI, 87–93%). The per-patient based pooled LR+ was 7.5 (95% CI, 4.55–12.37) and pooled LR– was 0.26 (95% CI, 0.14–0.48). The per-patient based pooled DOR was 36.7 (95% CI, 12.65–106.47). The area under the curves (AUC) was 0.93.

Among the studies with per-lymph node based data (Fig 3), the sensitivity of MRI ranged between 53% and 93%, and the specificity ranged between 91% and 99%. The pooled per-lymph node based sensitivity for MRI was 77% (95% CI, 72–82%), and the specificity was 98% (95% CI, 97–98%). The per-lymph node based pooled LR+ was 42.24 (95% CI, 11.28–158.13) and pooled LR– was 0.21 (95% CI, 0.09–0.47). The per-lymph node based pooled DOR was 212.35 (95% CI, 28.23–1597.46). The AUC was 0.98.

However, there was significant heterogeneity among included studies. The I^2 indices were higher than 61.9%, and the χ^2 statistics were significantly higher in per-patient and per-lymph node based pooled sensitivity, specificity, LR+, LR–, and DOR.

Meta-analysis of quantitative short time inversion recovery and diffusion-weighted imaging

For studies with per-patient based data, the diagnostic accuracy of quantitative STIR was reported in four studies,[38,39,44,45] and the diagnostic accuracy of quantitative DWI was reported in three studies.[44–46] We performed meta-analysis to compare the diagnostic value between quantitative STIR and DWI. Table 4 shows detailed results of the meta-analysis.

Table 1 Principle characteristics of included studies

Study	Year	Country	Patients(n)	Mean age (years)	Gender (M/F)	Study design	Patient enrollment	Histology Adeno/ squamos/other	N stage N0/N1/N2/N3	Data type	Reference test
Ohno *et al.*[38]	2004	Japan	110	64 (36–82)	68/42	prospective	consecutive	85/18/7	ND	Per-patient based Per-lymph node based	Pathological analysis (mediastinoscopy or thoracotomy)
Ohno *et al.*[39]	2007	Japan	115	68 (35–81)	59/56	prospective	consecutive	96/13/6	72/32. 10/1	Per-patient based Per-lymph node based	Pathological analysis (mediastinoscopy or thoracotomy)
Kim *et al.*[40]	2008	Korea	113	61 (34–82)	91/22	prospective	consecutive	58/41/14	62/23/24/4	Per-lymph node based	Pathological analysis (mediastinoscopy or thoracotomy)
Hasegawa *et al.*[41]	2008	Japan	42	66 (41–83)	30/12	prospective	consecutive	ND	34/3/5/0	Per-patient based	Pathological analysis (thoracotomy)
Nomori *et al.*[42]	2008	Japan	88	70 (38–82)	47/41	prospective	ND	67/18/3	71/9/8/0	Per-lymph node based	Pathological analysis (thoracotomy)
Yi *et al.*[43]	2008	Korea	165	61 (34–82)	125/40	prospective	consecutive	86/59/20	79/26/33/12	Per-patient based	Pathological analysis (mediastinoscopy or thoracotomy or PCNA)
Nakayama *et al.*[44]	2010	Japan	70	68 (48–82)	38/32	retrospective	ND	52/18/0	54/9/7/0	Per-patient based	Pathological analysis (thoracotomy)
Ohno *et al.*[45]	2011	Japan	250	73 (61–83)	136/114	prospective	consecutive	218/23/9	157/72/16/5	Per-patient based	Pathological analysis (mediastinoscopy or thoracotomy)
Usuda *et al.*[46]	2011	Japan	63	68 (38–81)	41/22	ND	ND	42/19/2	41/11/11/0	Per-patient based Per-lymph node based	Pathological analysis (thoracotomy)

M/F, male/female, ND, not documented; PCNA, percutaneous needle aspiration biopsy.

For quantitative STIR, the pooled sensitivity was 84% (95% CI, 78–89%), and the specificity was 91% (95% CI, 87–94%). The pooled LR+ was 8.44 (95% CI, 6.05–11.78) and pooled LR– was 0.18 (95% CI, 0.08–0.44). The pooled DOR was 56.29 (31.92–99.24).

For quantitative DWI, the pooled sensitivity was 69% (95% CI, 61–77%), and the specificity was 93% (95% CI, 89–96%). The pooled LR+ was 8.36 (95% CI, 5.05–13.83) and pooled LR– was 0.36 (95% CI, 0.26–0.5). The pooled DOR was 27.2 (14.62-50.60). The pooled DOR estimate for STIR was greater than for DWI.

Discussion

Accurate staging is of vital importance for patients with NSCLC in order to choose the best treatment, such as to avoid unnecessary surgery because of lymph node involvement in the mediastinum, or to guide target volume delineation for patients treated with radiotherapy. For N staging in NSCLC, MRI is one of the noninvasive modalities with high sensitivity and specificity. Compared with PET/CT, MRI is more accessible, less expensive, and there is no risk of radiation exposure. In addition, functional images such as DWI can also be obtained.[42] To our knowledge, there has been no systematic review and meta-analysis to evaluate the diagnostic performance of MRI for N staging in patents with NSCLC.

After a thorough data search and rigid study selection, we included nine studies evaluating the diagnostic value of MRI for lymph node staging of NSCLC. The methodological quality was moderate to good. Without regard to the MRI technique performed in each study, pooled sensitivity, specificity, LR+, LR–, and DOR for per-patient based analysis (7 studies) were 74%, 90%, 7.5, 0.26, and 36.7, respectively, and for per-lymph node based analysis (5 studies) were 77%, 98%, 42.24, 0.21, and 212.35, respectively. MRI showed a high specificity for lymph node staging in NSCLC.

However, there was significant heterogeneity among included studies, which may be attributed to many factors, such as different patient spectrums and MRI techniques. Considering the complexity of MRI, we hypothesized that different pulse sequences (such as STIR or DWI) and diagnostic criteria (quantitative or qualitative) used in studies might be the culprit. Therefore, we extracted data from per-patient based studies that evaluated diagnostic performance of quantitative STIR and DWI and conducted meta-analysis to compare their diagnostic value.

The DOR of a test obtained with different combinations of sensitivity and specificity can be used as a single summary measure and is the ratio of the odds of positivity in disease relative to the odds of positivity in the non-diseased.[36,49] In our meta-analysis, the pooled DOR for quantitative STIR and DWI was 56.29 and 27.20, respectively. The DOR estimate for STIR was greater than that for DWI. However, such evidence

Table 2 Characteristics of MRI of included studies

Study	Magnet	Pulse sequences	Diagnostic criteria
Ohno *et al*. 2004[38]	1.5-T superconducting magnet	Transverse ECG and respiratory-triggered STIR TSE	Quantitative: LSR ≥ 0.6. Qualitative: signal intensity of lymph node was greater than that of muscle.
Ohno *et al*. 2007[39]	1.5-T superconducting magnet	Axial and coronal STIR TSE	Quantitative: LSR ≥ 0.6.
Kim *et al*.[40]	3-T superconducting magnet	Breath-hold T1-weighted TFE sequence Breath-hold cardiac-gated T2-weighted TSE (TIBB)	Quantitative: LTR ≥ 0.84. Qualitative: nodal morphologic characteristics (eccentric cortical thickening or obliteration of the fatty hilum of lymph node); lymph node size.
Hasegawa *et al*.[41]	1.5-T superconducting magnet	Transverse non-breath-hold DWI (STIR EPI) Transverse electrocardiographically and respiratory-triggered T2-weighted sequence	Qualitative: lymph node metastasis was defined as a focus of low signal intensity on DWI with a visible lymph node on corresponding T2-weighted image.
Nomori *et al*.[42]	1.5-T superconducting magnet	Coronal T1-weighted sequence Coronal and axial T2-weighted sequence Coronal and axial STIR sequence Transverse DWI (EPI)	Quantitative: ADC_{LN}-min ≤ 1.6×10^{-3} mm^2/s.
Yi *et al*.[43]	3-T superconducting magnet	Breath-hold T1-weighted TFE sequence Breath-hold cardiac-gated T2-weighted TSE (TIBB)	Qualitative: nodal morphologic characteristics (eccentric cortical thickening or obliteration of the fatty hilum of lymph node).
Nakayama *et al*.[44]	1.5-T superconducting magnet	Transverse T1-weighted and T2-weighted sequences Transverse DWI (HASTE) Transverse breath-hold STIR TSE	Quantitative STIR: LSR ≥ 0.354. Quantitative DWI: $ADC_{LN} \leq 1.54 \times 10^{-3}$ mm^2/s. ADC_{LC}-$ADC_{LN} \leq 0.24 \times 10^{-3}$ mm^2/s.
Ohno *et al*. 2011[45]	1.5-T superconducting magnet	Axial and coronal breath-hold STIR TSE Three axes (axial, sagittal, and coronal) DWI (STIR EPI)	Quantitative STIR: LSR ≥ 0.6. LMR ≥ 1.4. Quantitative DWI: $ADC_{LN} \leq 2.5 \times 10^{-3}$ mm^2/s. Qualitative STIR or DWI: signal intensity of lymph node was greater than that of muscle.
Usuda *et al*.[46]	1.5-T superconducting magnet	Coronal T1-weighted SE Coronal and axial T2-weighted FSE Respiratory triggered DWI (SS EPI with SPAIR)	Quantitative DWI: $ADC_{LN} \leq 1.7 \times 10^{-3}$ mm^2/s.

ADC_{LC}, apparent diffusion coefficient value of lung cancer; ADC_{LN}, ADC value of lymph node; DWI, diffusion-weighted imaging; ECG, electrocardiogram; EPI, echo-planar imaging; FS, fat suppression; FSE, fast spin-echo; HASTE, half-Fourier acquisition single-shot turbo spin echo; LMR, lymph node to muscle ratio of signal intensity; LSR, lymph node to saline ratio of signal intensity; LTR, lymph node to tumor ratio of signal intensity; MRI, magnetic resonance imaging; SPAIR, spectral presaturation attenuated inversion recovery; SS, single shot; STIR, short time inversion recovery; TFE, turbo field-echo; TIBB, triple-in-version black-blood; TSE, turbo spin echo.

Table 3 Evaluation of quality of included studies using QUADAS

	Ohno 2004	Ohno 2007	Kim	Hasegawa	Nomori	Yi	Nakayama	Ohno 2011	Usuda
Representative spectrum?	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	No
Acceptable reference standard?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Acceptable delay between tests?	Unclear	Unclear	Yes	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Partial verification avoided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Differential verification avoided?	No	No	No	Yes	Yes	No	Unclear	No	No
Incorporation avoided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reference standard results blinded?	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Index test results blinded?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Relevant clinical information?	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Yes
Uninterpretable results reported?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Withdrawals explained?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

QUADAS, Quality Assessment of Diagnostic Accuracy Studies.

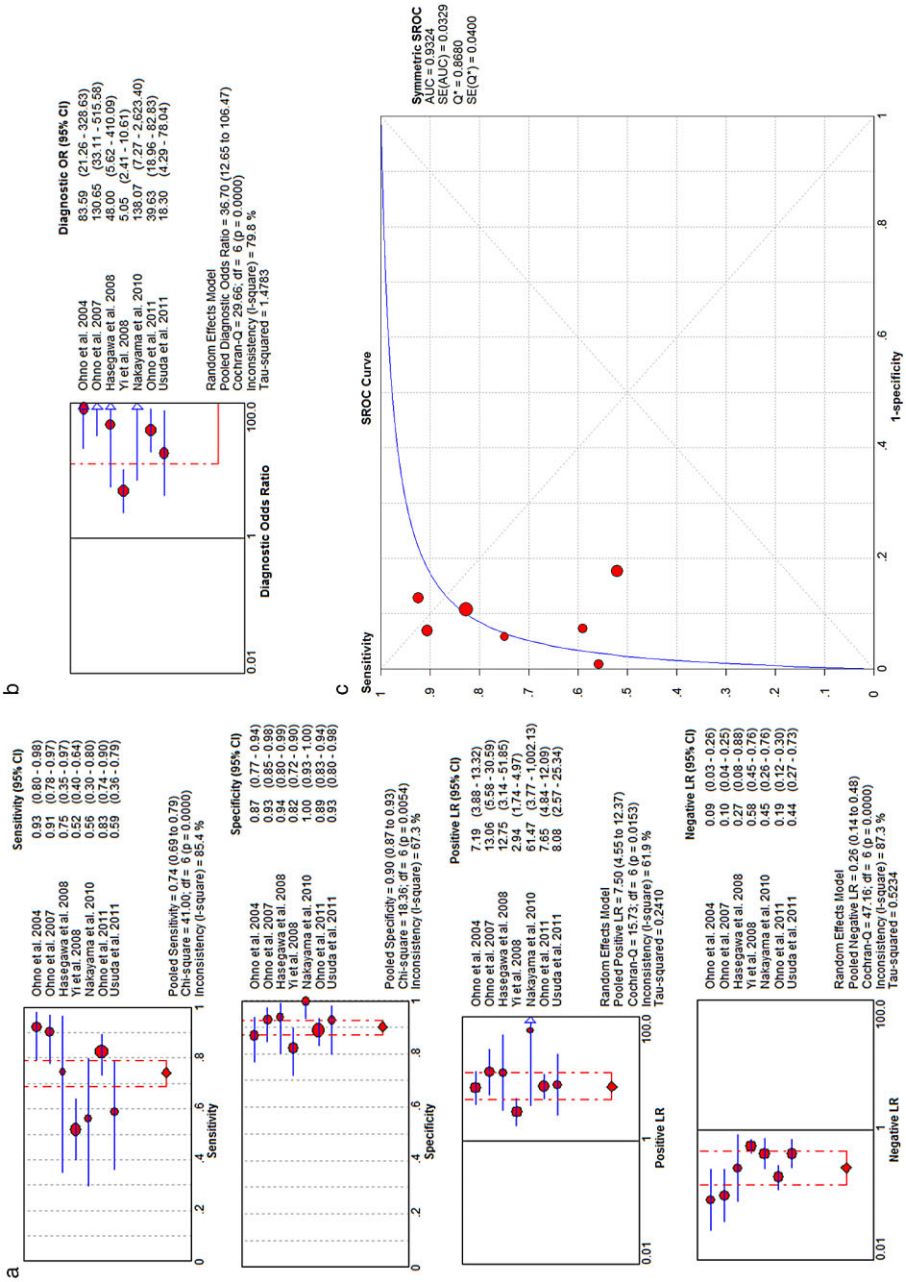

Figure 2 (a) Per-patient based pooled data (sensitivity, specificity, positive likelihood ratio [LR+], negative likelihood ratio [LR−]) and (b) DOR (c) SROC curve.

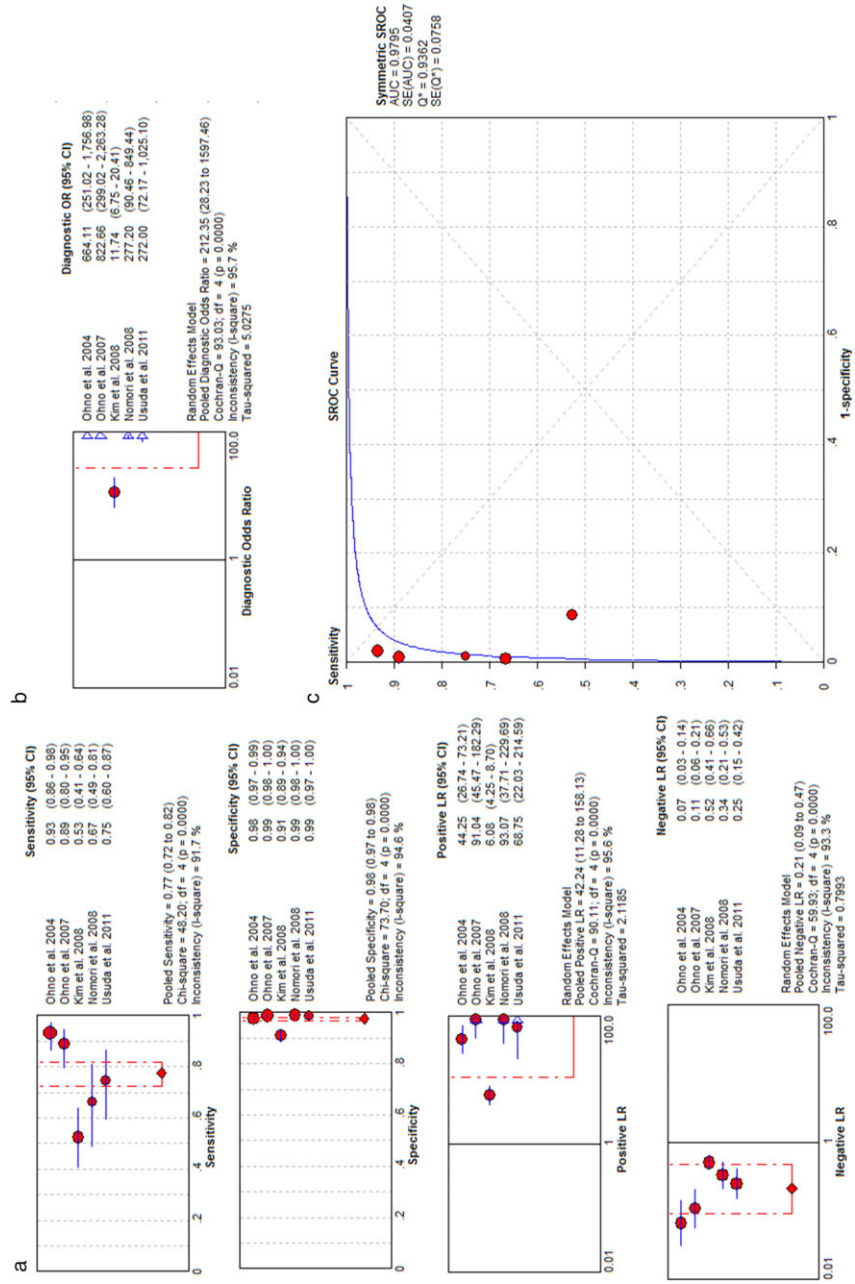

Figure 3 (a) Per-lymph node based pooled data (sensitivity, specificity, LR+, LR–); (b) diagnosis odds ratio (DOR); (c) summary receiver operating characteristic (SROC) curve.

Table 4 Diagnostic accuracy of quantitative STIR and DWI in evaluation of N-staging in NSCLC patients (per-patients basis)

MRI method	No. of patients	Pooled sensitivity	Pooled specificity	LR+	LR–	DOR	Heterogeneity
STIR	545	0.84 (0.78–0.89)	0.91 (0.87–0.94)	8.44 (6.05–11.78)	0.18 (0,08–0.44)	56.29 (31.92–99.24)	$0\% \le I^2 \le 84.9\%$
DWI	383	0.69 (0.61–0.77)	0.93 (0.89–0.96)	8.36 (5.05–13.83)	0.36 (0.26–0.5)	27.2 (14.64–50.60)	$0\% \le I^2 \le 78.4\%$

Data in parentheses are 95% confidence intervals. DOR, diagnosis odds ratio; DWI, diffusion-weighted imaging; LR–, negative likelihood ratio; LR+, positive likelihood ratio; NSCLC, non-small cell lung cancer; STIR, short time inversion recovery.

was based on indirect comparison, with only 545 and 383 patients in the quantitative STIR and DWI groups, respectively. In addition, a different cut-off was used in each study, which may cause a threshold effect. Therefore, large, direct, and prospective comparative studies are needed to confirm the superiority of STIR or DWI.

A recent meta-analysis compared DWI with F-FDG PET/CT for pre-operative staging of mediastinal and hilar lymph nodes in NSCLC patients.[18] The authors concluded that DWI has a high specificity for N staging of NSCLC.[36] While our study also showed high specificity of DWI, the pooled LR+ and DOR of DWI was lower; as our study analyzed different studies, we confined analysis to quantitative DWI, which may have different diagnostic power compared with qualitative DWI. Such evidence indicates that a standard diagnostic criterion is needed for each MRI technique to distinguish metastatic and non-metastatic lymph nodes.

Our study has several limitations. First, only nine studies were included, seven studies for per-patient analysis, and five studies for per-lymph node analysis. Second, all of the included studies were performed in Japan and Korea, and we confined the literature search to English language articles, leading to a publication bias. Third, differential verification bias could not be avoided in six studies and is unclear in one study. Fourth, the studies used different MRI pulse sequences and different diagnostic criteria, which lead to different sensitivity and specificity. Fifth, only four and three studies were used to pool diagnostic performance of STIR and DWI, respectively.

Conclusions

Despite several limitations, our study has confirmed that MRI has high specificity for N staging in NSCLCs, which is helpful for developing optimal clinical therapeutic strategies. In addition, quantitative STIR has greater DOR than quantitative DWI in the lymph node staging of NSCLC. However, before wide application of MRI for N staging in clinical practice, the best protocols of pulse sequences and consistent diagnostic criteria should be identified. Larger prospective studies are warranted to compare quantitative or qualitative STIR and DWI.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 81272501).

Disclosure

No authors report any conflict of interest.

References

1 Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. (Published erratum appears in *CA Cancer J Clin* 2014; **64:** 364) *CA Cancer J Clin* 2014; **64**: 9–29.
2 Howlader N, Noone A, Krapcho M. *SEER Cancer Statistics Review, 1975–2009 (Vintage 2009 Populations)*. National Cancer Institute, Bethesda 2012.
3 Edge S, Byrd D, Compton C, Fritz AG, Greene FL, Trotti A (eds). *AJCC Cancer Staging Manual, 7th Ed*. Springer, New York 2010.
4 Martins RG, D'Amico TA, Loo BW, Jr *et al*. The management of patients with stage IIIA non-small cell lung cancer with N2 mediastinal node involvement. *J Natl Compr Canc Netw* 2012; **10**: 599–613.
5 Fernandes AT, Shen J, Finlay J *et al*. Elective nodal irradiation (ENI) vs. involved field radiotherapy (IFRT) for locally advanced non-small cell lung cancer (NSCLC): a comparative analysis of toxicities and clinical outcomes. *Radiother Oncol* 2010; **95**: 178–84.
6 Liberman M, Sampalis J, Duranceau A, Thiffault V, Hadjeres R, Ferraro P. Endosonographic mediastinal lymph node staging of lung cancer. *Chest* 2014; **146**: 389–97.
7 Navani N, Spiro SG, Janes SM. Mediastinal staging of NSCLC with endoscopic and endobronchial ultrasound. *Nat Rev Clin Oncol* 2009; **6**: 278–86.
8 Walker CM, Chung JH, Abbott GF *et al*. Mediastinal lymph node staging: from noninvasive to surgical. *AJR Am J Roentgenol* 2012; **199**: W54–64.
9 Paul NS, Ley S, Metser U. Optimal imaging protocols for lung cancer staging: CT, PET, MR imaging, and the role of imaging. *Radiol Clin North Am* 2012; **50**: 935–49.
10 Boiselle PM, Patz EF, Jr, Vining DJ, Weissleder R, Shepard JA, McLoud TC. Imaging of mediastinal lymph nodes: CT, MR, and FDG PET. *Radiographics* 1998; **18**: 1061–9.

11 Shim SS, Lee KS, Kim BT *et al*. Non-small cell lung cancer: prospective comparison of integrated FDG PET/CT and CT alone for preoperative staging. *Radiology* 2005; **236**: 1011–9.
12 Antoch G, Stattaus J, Nemat AT *et al*. Non-small cell lung cancer: dual-modality PET/CT in preoperative staging. *Radiology* 2003; **229**: 526–33.
13 Halpern BS, Schiepers C, Weber WA *et al*. Presurgical staging of non-small cell lung cancer: positron emission tomography, integrated positron emission tomography/CT, and software image fusion. *Chest* 2005; **128**: 2289–97.
14 Birim O, Kappetein AP, Stijnen T, Bogers AJ. Meta-analysis of positron emission tomographic and computed tomographic imaging in detecting mediastinal lymph node metastases in nonsmall cell lung cancer. *Ann Thorac Surg* 2005; **79**: 375–82.
15 Levitt RG, Glazer HS, Roper CL, Lee JK, Murphy WA. Magnetic resonance imaging of mediastinal and hilar masses: comparison with CT. *AJR Am J Roentgenol* 1985; **145**: 9–14.
16 Martini N, Heelan R, Westcott J *et al*. Comparative merits of conventional, computed tomographic, and magnetic resonance imaging in assessing mediastinal involvement in surgically confirmed lung carcinoma. *J Thorac Cardiovasc Surg* 1985; **90**: 639–48.
17 Patterson GA, Ginsberg RJ, Poon PY *et al*. A prospective evaluation of magnetic resonance imaging, computed tomography, and mediastinoscopy in the preoperative assessment of mediastinal node status in bronchogenic carcinoma. *J Thorac Cardiovasc Surg* 1987; **94**: 679–84.
18 Laurent F, Drouillard J, Dorcier F *et al*. Bronchogenic carcinoma staging: CT versus MR imaging. Assessment with surgery. *Eur J Cardiothorac Surg* 1988; **2**: 31–6.
19 Georgian D, Rice TW, Mehta AC, Wiedemann HP, Stoller KJ, O'Donovan PB. Intrathoracic lymph node evaluation by CT and MRI with histopathologic correlation in non-small cell bronchogenic carcinoma. *Clin Imaging* 1990; **14**: 35–40.
20 Stiglbauer R, Schurawitzki H, Klepetko W *et al*. Contrast-enhanced MRI for the staging of bronchogenic carcinoma: comparison with CT and histopathologic staging–preliminary results. *Clin Radiol* 1991; **44**: 293–8.
21 Mayr B, Lenhard M, Fink U, Heywang-Köbrunner SH, Sunder-Plassmann L, Permanetter W. Preoperative evaluation of bronchogenic carcinoma: value of MR in T- and N-staging. *Eur J Radiol* 1992; **14**: 245–51.
22 Webb WR, Gatsonis C, Zerhouni EA *et al*. CT and MR imaging in staging non-small cell bronchogenic carcinoma: report of the Radiologic Diagnostic Oncology Group. *Radiology* 1991; **178**: 705–13.
23 Manfredi R, Pirronti T, Bonomo L, Marano P. Accuracy of computed tomography and magnetic resonance imaging in staging bronchogenic carcinoma. *MAGMA* 1996; **4**: 257–62.
24 Crisci R, Di Cesare E, Lupattelli L, Coloni GF. MR study of N2 disease in lung cancer: contrast-enhanced method using gadolinium-DTPA. *Eur J Cardiothorac Surg* 1997; **11**: 214–7.
25 Kernstine KH, Stanford W, Mullan BF *et al*. PET, CT, and MRI with Combidex for mediastinal staging in non-small cell lung carcinoma. *Ann Thorac Surg* 1999; **68**: 1022–8.
26 Nguyen BC, Stanford W, Thompson BH *et al*. Multicenter clinical trial of ultrasmall superparamagnetic iron oxide in the evaluation of mediastinal lymph nodes in patients with primary lung carcinoma. *J Magn Reson Imaging* 1999; **10**: 468–73.
27 Takenaka D, Ohno Y, Hatabu H *et al*. Differentiation of metastatic versus non-metastatic mediastinal lymph nodes in patients with non-small cell lung cancer using respiratory-triggered short inversion time inversion recovery (STIR) turbo spin-echo MR imaging. *Eur J Radiol* 2002; **44**: 216–24.
28 Hasegawa I, Eguchi K, Kohda E *et al*. Pulmonary hilar lymph nodes in lung cancer: assessment with 3D-dynamic contrast-enhanced MR imaging. *Eur J Radiol* 2003; **45**: 129–34.
29 Gjurchinov D, Stojanovska-Nojkova J, Grunevski M, Nikolov V, Spasovski G. Radiological and "imaging" methods in TNM classification of non-small-cell lung cancer. *Prilozi* 2007; **28**: 155–67.
30 Plathow C, Aschoff P, Lichy MP *et al*. Positron emission tomography/computed tomography and whole-body magnetic resonance imaging in staging of advanced nonsmall cell lung cancer – initial results. *Invest Radiol* 2008; **43**: 290–7.
31 Morikawa M, Demura Y, Ishizaki T *et al*. The effectiveness of 18F-FDG PET/CT combined with STIR MRI for diagnosing nodal involvement in the thorax. *J Nucl Med* 2009; **50**: 81–7.
32 Chen W, Jian W, Li HT *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* 2010; **28**: 613–20.
33 Pauls S, Schmidt SA, Juchems MS *et al*. Diffusion-weighted MR imaging in comparison to integrated [^{18}F]-FDG PET/CT for N-staging in patients with lung cancer. *Eur J Radiol* 2012; **81**: 178–82.
34 Schwenzer NF, Schraml C, Müller M *et al*. Pulmonary lesion assessment: comparison of whole-body hybrid MR/PET and PET/CT imaging–pilot study. *Radiology* 2012; **264**: 551–8.
35 Sommer G, Wiese M, Winter L *et al*. Preoperative staging of non-small-cell lung cancer: comparison of whole-body diffusion-weighted magnetic resonance imaging and 18F-fluorodeoxyglucose-positron emission tomography/computed tomography. *Eur Radiol* 2012; **22**: 2859–67.
36 Wu LM, Xu JR, Gu HY *et al*. 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* 2012; **178**: 304–14.
37 Kim YN, Yi CA, Lee KS *et al*. A proposal for combined MRI and PET/CT interpretation criteria for preoperative nodal staging in non-small-cell lung cancer. *Eur Radiol* 2012; **22**: 1537–46.
38 Ohno Y, Hatabu H, Takenaka D *et al*. Metastases in mediastinal and hilar lymph nodes in patients with non-small

cell lung cancer: quantitative and qualitative assessment with STIR turbo spin-echo MR imaging. *Radiology* 2004; **231**: 872–9.
39 Ohno Y, Koyama H, Nogami M *et al*. STIR turbo SE MR imaging vs. coregistered FDG-PET/CT: quantitative and qualitative assessment of N-stage in non-small-cell lung cancer patients. *J Magn Reson Imaging* 2007; **26**: 1071–80.
40 Kim HY, Yi CA, Lee KS *et al*. Nodal metastasis in non-small cell lung cancer: accuracy of 3.0-T MR imaging. *Radiology* 2008; **246**: 596–604.
41 Hasegawa I, Boiselle PM, Kuwabara K, Sawafuji M, Sugiura H. Mediastinal lymph nodes in patients with non-small cell lung cancer: preliminary experience with diffusion-weighted MR imaging. *J Thorac Imaging* 2008; **23**: 157–61.
42 Nomori H, Mori T, Ikeda K *et al*. 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* 2008; **135**: 816–22.
43 Yi CA, Shin KM, Lee KS *et al*. Non-small cell lung cancer staging: efficacy comparison of integrated PET/CT versus 3.0-T whole-body MR imaging. *Radiology* 2008; **248**: 632–42.
44 Nakayama J, Miyasaka K, Omatsu T *et al*. 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* 2010; **34**: 1–8.
45 Ohno Y, Koyama H, Yoshikawa T *et al*. 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* 2011; **261**: 605–15.
46 Usuda K, Zhao XT, Sagawa M *et al*. Diffusion-weighted imaging is superior to positron emission tomography in the detection and nodal assessment of lung cancers. *Ann Thorac Surg* 2011; **91**: 1689–95.
47 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; **3**: 25.
48 Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C (eds). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0*. The Cochrane Collaboration, 2009. [Cited: 22 Nov 2014.] Available from URL: http://srdta.cochrane.org/
49 Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994; **271**: 389–91.

