Medicine®

SYSTEMATIC REVIEW AND META-ANALYSIS

OPEN

Diagnostic Performance of Diffusion-weighted Magnetic Resonance Imaging in Bone Malignancy

Evidence From a Meta-Analysis

Li-Peng Liu, PhD, Long-Biao Cui, PhD, Xin-Xin Zhang, PhD, Jing Cao, PhD, Ning Chang, PhD, Xing Tang, PhD, Shun Qi, PhD, Xiao-Liang Zhang, PhD, Hong Yin, MD, PhD, and Jian Zhang, MD, PhD

Abstract: Current state-of-the-art nuclear medicine imaging methods (such as PET/CT or bone scintigraphy) may have insufficient sensitivity for predicting bone tumor, and substantial exposure to ionizing radiation is associated with the risk of secondary cancer development. Diffusion-weighted MRI (DW-MRI) is radiation free and requires no intravenous contrast media, and hence is more suitable for population groups that are vulnerable to ionizing radiation and/or impaired renal functions. This meta-analysis was conducted to investigate whether whole-body DW-MRI is a viable means in differentiating bone malignancy.

Medline and Embase databases were searched from their inception to May 2015 without language restriction for studies evaluating DW-MRI for detection of bone lesions. Methodological quality was assessed by the quality assessment of diagnostic studies (QUADAS-2) instrument. Sensitivities, specificities, diagnostic odds ratio (DOR), and areas under the curve (AUC) were used as measures of the diagnostic accuracy. We combined the effects by using the random-effects mode. Potential threshold effects and publication bias were investigated.

We included data from 32 studies with 1507 patients. The pooled sensitivity, specificity, and AUC were 0.95 (95% CI, 0.90–0.97), 0.92 (95% CI, 0.88–0.95), and 0.98 on a per-patient basis, and they were 0.91 (95% CI, 0.87–0.94), 0.94 (95% CI, 0.90–0.96), and 0.97 on a per-lesion basis. In subgroup analysis, there is no statistical significance found in the sensitivity and specificity of using DWI only and DWI combined with other morphological or functional imaging sequence in both basis ($P > 0.05$). A b value of 750 to 1000 s/mm^2 enables higher AUC and DOR for whole-body imaging purpose when compared with other values in both basis either ($P < 0.01$). The ROC space did not show a curvilinear trend of points and a threshold effect was not observed. According to the Deek's plots, there was no publication bias on both basis.

Our results support the use of DWI as an effective means for distinguishing malignant bone lesions; however, various imaging parameters need to be standardized prior to its broad use in clinical practice.

(*Medicine* 94(45):e1998)

Abbreviations: ADC = apparent diffusion coefficient, ASCO = American Society of Clinical Oncology, AUC = areas under the curve, BEIR = Biological Effect of Ionizing Radiation, CI = confidence interval, CMS = Centers for Medicare and Medicaid Services, DOR = diagnostic odds ratio, DW-MRI = diffusion-weighted magnetic resonance imaging, ESMO = European Society for Medical Oncology, FDG-[18F]2 = fluoro-2-D-glucose, FN = false negative, FP = false positive, LR^+ and LR^- = positive likelihood ratio and negative likelihood ratio, NCCN = National Comprehensive Cancer Network, PET/CT = positron emission tomography/computed tomography, PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses, QUADAS = Quality Assessment of Diagnostic Accuracy Studies, SROC = summary receiver-operator characteristic, Tc-99m MDP BS = technetium-99m methyldiphosphonate bone scintigraphy, TN = true negatives, TP = true positives.

INTRODUCTION

Malignant tumors are associated with poor clinical outcomes and high morbidity and mortality,[1–4] as compared with benign tumors. Hence, an effective means for differentiating between malignant tumors and benign tumors is crucial for accurate diagnosis. Whole-body technetium-99m methyldiphosphonate (Tc-99m MDP) bone scintigraphy (BS) is one of the most commonly practiced methods for suspected bone lesions, especially for patients with pain symptom in follow-up visits, and it remains a reference for oncologists.[5] However, the main limitation of BS is the fact that detection for new lesion becomes difficult for patients already exhibiting elevated Tc-99m MDP uptake, which subsequently affects its accuracy in predicting bone tumor. Magnetic resonance imaging (MRI) and positron emission tomography (PET) have shown potential in bone tumor diagnosis, but their replacement of BS is still debatable.[6]

PET/CT has combined benefit of PET's sensitivity and CT's anatomical information, and metabolic tracers such as

Editor: Wael Alkhiary.
Received: June 1, 2015; revised: October 14, 2015; accepted: October 16, 2015.
From the Department of Respiratory Medicine (L-PL, X-XZ, JC, NC, JZ); Department of Radiology, Xijing Hospital, Fourth Military Medical University, Xi'an, China (L-BC, XT, SQ, HY); and Department of Radiology, University of California, San Francisco, California (X-LZ).
Correspondence: Jian Zhang, Department of Respiratory Medicine, Xijing Hospital, Fourth Military Medical University, 127 Changle West Road, Xi'an 710032, China (e-mail: zhangjian197011@yahoo.com). Hong Yin, Department of Radiology, Xijing Hospital, Fourth Military Medical University, 127 Changle West Road, Xi'an 710032, China (e-mail: yinhong@fmmu.edu.cn).
Supplemental Digital Content is available for this article.
L-PL, L-BC, and X-XZ contributed equally to this work as first author.
Authors' contributions conception and design: HY, JZ
Development of methodology: L-PL.
Acquisition of data: L-PL, L-BC, X-XZ, JC, NC.
Writing, review, and/or revision of the manuscript: HY, JZ, L-PL, L-BC.
Analysis and interpretation of data (eg, statistical analysis, biostatistics, computational analysis): L-PL.
Administrative, technical, or material support (ie, reporting or organizing data, constructing databases): XT, SQ.
Study supervision: HY, JZ.
This work was supported by the Shaanxi Research Project (2010K14-01-01).
The authors have no conflicts of interest to disclose.

ISSN: 0025-7974
DOI: 10.1097/MD.0000000000001998

[18F]2-fluoro-2-D-glucose (FDG) can show the elevated glucose uptake and cellular metabolism within tumors. PET/CT has superior sensitivity in detecting bone metastases than other imaging modalities,[7] however is insensitive in detecting bone marrow involvement.[8] Recently, European guidelines addressing PET/CT in bony tumors has concluded that ''the role of PET-CT in monitoring bone lesions has been reported in a few small studies and appears potentially promising; however, prospective trials are needed to establish its true clinical utility.''[9] The major pitfall for PET/CT is associated with ionizing radiation exposure. With the recent decision to end the National Oncology PET Registry, use of PET/CT for routine surveillance is now clearly not recommended by Centers for Medicare and Medicaid Services (CMS).[10] American Society of Clinical Oncology (ASCO), the European Union of Urology, the European Society for Medical Oncology (ESMO),[11] and NCCN have all declined to include surveillance PET examinations in disease-specific guidelines. Moreover, ionizing radiation can also be produced by bone scintigraphy which is the same as the PET/CT. As for the amount of ionizing radiation, US annual per capita radiation dose increased from 0.1 mSv in 1980 to 0.77 mSv in 2006 from the source of nuclear medicine.[12] According to the BEIR report VII, exposure to ionizing radiation causes roughly a tripling effect in lifetime cancer risk by comparison with a person without exposure above the age of 30 years.[13]

Diffusion-weighted imaging (DWI) is a technique that probes the level of water molecule diffusion within the microstructures in tissues, and is sensitive to local pathological alterations. Over years, DWI has found a broad range of applications both in neuro imaging and in body imaging; moreover, it has gained much attention in tumor imaging due to both its outstanding sensitivity and specificity as well as the absence of contrast media administration, which is important for patients with impaired renal functions. The level of diffusion is controlled by the diffusion-sensitizing coefficient so called *b*-value, and diffusion acquisition at two distinctive *b*-values (zero or nonzero) allows the derivation of the apparent diffusion coefficient (ADC, in the unit of s/mm^2). ADC value is a quantitative measure of water movement in tumors: low ADC values indicate an abundance of cell membranes, whereas high ADC values are indicative of cellular regions. Hence, DWI may be either qualitatively inspected or quantitatively assessed based on calculated ADC values.[14] Whole-body DW-MRI has recently become practical due to technological advances, and it emerged as a promising bone marrow assessment tool for detection of both primary cancer or distant metastasis of bone.[15,16] The added diagnostic value of DWI has been reported in several studies; there are consistent findings showing that it could have a comparable or better performance in diagnosing bone tumors, as compared with BS[17–22] or PET/CT[23–25].

Since various studies predicting the accuracy of DWI in detecting bone tumors have been published, results of these studies are drastically diverse because of the differing DWI protocol used. Here, we performed an updated meta-analysis to investigate the diagnostic value of DW-MRI as a standalone method in bone lesions screening. Moreover, we intended to compare the average adjusted accuracies of DWI between different DWI sequences, analysis methods, *b* values, and covariate that may affect the effectiveness of modalities.

METHODS

We did a meta-analysis in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines [26] and the guidelines described in the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy.

Search Strategy and Selection Criteria

We searched for the relevant studies (Table S1, http://links.lww.com/MD/A514) in the online database of EMBASE, PubMed from the date of their inception up to May 2015 with the assistance of a librarian. No language restriction was placed. The reference in all the retrieved articles was also searched for any additional relevant studies. A radiologist and an oncologist were asked to look through these literatures and assess their eligibility for analysis.

The inclusion criteria included: studies that assessed the sensitivity, specificity, and other metrics assessing the diagnostic performance of DWI, among which systematic reviews and meta-analyses were used only as a source of references; studies that validated the performance of DWI in cancer diagnosis and showed that all participants had the reference tests; studies that assessed primary bone tumors or bone metastasis; and studies based on which the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) were able to be calculated on the basis of sensitivity and specificity in respective publications. Conference abstracts were also included when they contained relevant published data or relevant unpublished data if could be traced from the authors. We excluded all studies that could be classified as narrative reviews, letters, editorials, comments, and case reports, and surveillance of the response of chemoradiotherapy in patients with cancers. A total of 32 studies were finalized, any disagreement between them was resolved by discussing with a third party. The inclusion of all the studies based on the above criteria was done in 2 stages: in the first stage the inclusion was based on title and abstract; and in the second stage, the full texts were considered. The literature flow diagram is shown in the Appendix as PRISMA flowchart.

Quality Assessment

The quality of the selected studies and the potential bias were assessed using the prespecified QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies) guideline, including additional items as recommended by the Cochrane Collaboration.[27] This quality assessment procedure was independently performed by 2 pairs of reviewers and was checked by a fifth reviewer. Any disagreements were resolved by discussion involving all researchers when necessary. The reference standard was validated by a clinical review committee consisting of 3 researchers.

Data Extraction

Two reviewers independently extracted relevant data from the selected studies in a standard form, a third investigator checked the extracted data, and a fourth investigator arbitrated on discrepancies between the first 2 investigators. Any identified discrepancies were discussed and corrected. Two-by-two contingency tables were constructed based on the data published, summarizing TP, FP, TN, and FN on the basis of sensitivity and specificity in respective papers. Moreover, if various kinds of sequences (DWI sequence only vs. DWI combined with other functional imaging sequences) were available in same papers, we placed them in our study separately and made a subgroup analysis of each type in all studies. In the publications, either the number of patients or the number of

lesions was used for the statistical analyses; we conducted separate analyses for each category to avoid any potential inconsistency.

Statistical Analysis

A random effects model was performed for the primary meta-analysis using a nonlinear mixed model approach. The primary objective was to estimate the sensitivity and specificity, positive and negative likelihood ratios (LR^+ and LR^-), and diagnostic odds ratio (DOR) with 95% confidence intervals (CIs) of DW-MRI for the diagnosis of bone lesions. We assumed bivariate normal distributions for sensitivity and specificity and presented a forest plot. LR^+ and LR^- are metrics derived from the summarized sensitivity and specificity for assessing the discriminating ability of the imaging modality.[28,29] If the LR^+ is >5.0 and the LR^- is <0.2, then the test can both rule in and rule out the disease. DORs were calculated for the discriminating ability of the imaging methods. The value ranged from 0 to infinity, higher values indicate better discriminatory test performances. The receiver-operator characteristic (SROC) graph analyzes the pooled accuracy, and each data point comes from the different studies. SROC curve is then formed based on these points to form a smooth curve. The area under the SROC curve (AUC) was estimated for the diagnostic accuracy of each imaging method. An AUC that is >0.5 and closer to 1.0 implies better accuracy.

The presence of heterogeneity was assessed using a fixed-effect meta-regression and I^2 statistics.[30] I^2 over 0.50 indicates heterogeneous, while $P < 0.05$ was considered having heterogeneity in likelihood ratio χ^2 test. We assessed publication bias by Deeks' plots.[31] We analyzed data separately on a per-lesion and on a per-patient basis. We also performed separate analyses in which MR was performed with DWI only or DWI combined with other functional sequences. Furthermore, we performed separate analyses for the subset of *b* value (750–1000 s/mm^2 or other values). We also performed meta-analyses for within-study comparisons on the reference standard, factors related to study design (prospective or retrospective), consecutive enrollment and operation interval. All tests were 2-sided with a type I error of 0.05. All analyses were performed using the software StataSE version 12 (StataCorp) and MetaDisc (Version 1.4).

Ethics

All the data involved in this meta-analysis study were from sourced respective publications, which had their own ethic approval in accordance with the local ethic committee's guideline. Hence, no separate ethical committee approval is needed for this study.

RESULTS

Literature Searches

A total of 491 publications were reviewed. The filtering process for the publications is shown in the flowchart in Figure 1A, 459 publications were excluded after primary and subsequent reviewing. In the end, 32 papers involved 1507 patients, were included in this meta-analysis.[17–25,32–54] In addition, 8 papers[17,18,34,39,43,44,48,53] consisted of both per-lesion and per-patient based analyses, contributing additional 10 studies. Finally, 23 studies reported on a per-lesion basis and 19 studies on a per-patient basis met the inclusion criteria of our research. The detailed baseline characteristics are listed in Table 1.

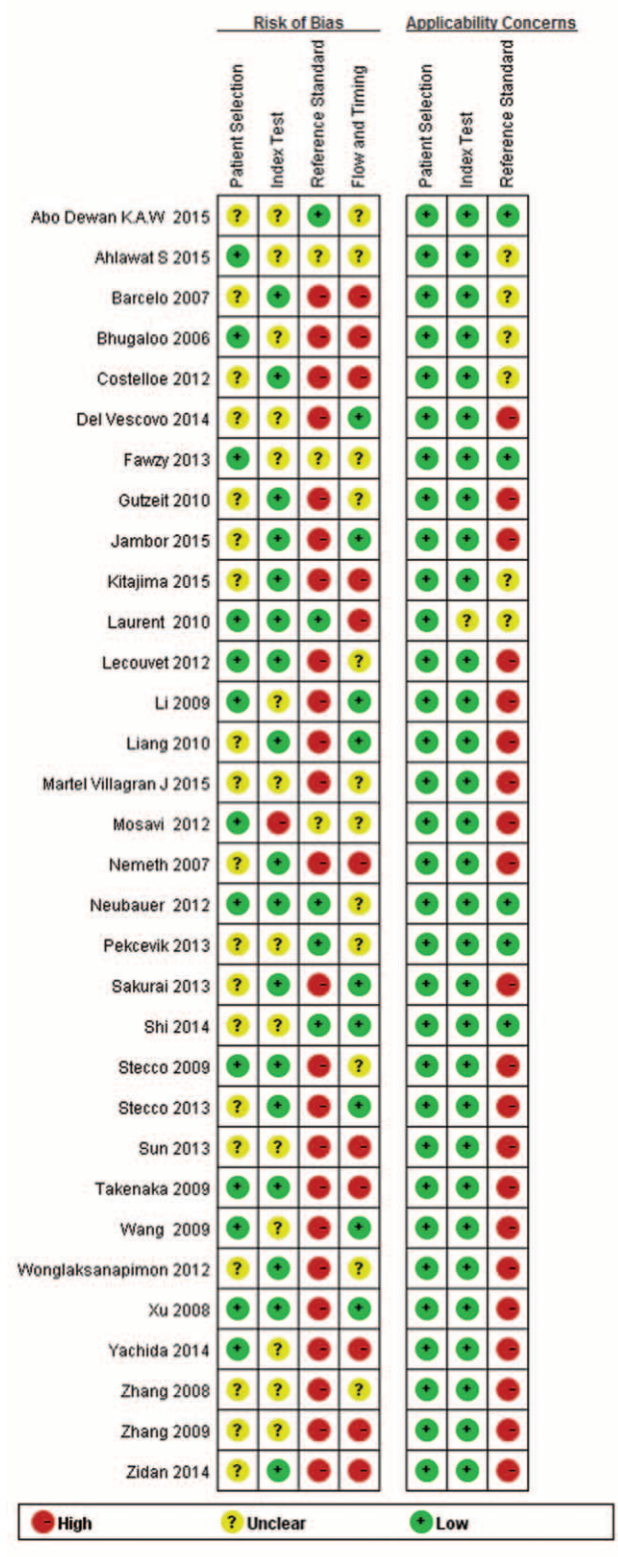

FIGURE 1. Summary of methodological quality of included studies on the basis of QUADAS-2 checklist for each study.

TABLE 1. Baseline Characteristics of the Included Studies

Study	Country	Patient Number	Man/ Women	Age	Design	Consecutive Enrollment	Malignancy	Magnetic Field Strength, T	*b* Value, s/mm^2	Modality	Reference	Analytical Method
Del Vescovo (17)	Italy 2014	17	6/11	57	Retrospective	ND	Multi	1.5	50, 500	DWI with other sequence	Imaging	Qualitative
Zidan (38)	Egypt 2014	56	31/25	46	Prospective	ND	Multi	1.5	0, 800	DWI with other sequence	Histology and follow-up	Qualitative
Yachida (37)	Japan 2014	27	27/0	72	Retrospective	Yes	Prostate cancer	1.5	0, 1000, 2000	DWI only	Imaging and follow-up	Quantitative
Sun (39)	China 2013	62	42/20	57.1	Prospective	ND	Multi	1.5	0, 800	DWI with other (DWIBS)	Imaging and follow-up	Qualitative
Stecco (18)	Italy 2013	23	ND	ND	ND	ND	Multi	1.5	1000	DWI with other (DWIBS)	Imaging	Qualitative
Sakurai (23)	Japan 2013	23	7/16	56	Prospective	ND	Thyroid	1.5	800	DWI with other (DWIBS)	Imaging and follow-up	Qualitative
Fawzy (40)	Egypt 2013	100	65/35	61	Prospective	Yes	Primary bone	1.5	500, 800	DWI with other sequence	Histology	Quantitative
Wonglaksanapimon (47)	Thailand 2012	22	8/14	54.6	Prospective	ND	Multi	3	400	DWI only	Histology or imaging studies	Quantitative
Neubauer (32)	Germany 2012	44	18/26	11	Retrospective	Yes	Primary bone	1.5 + 3	50, 800	DWI only	Histology	Quantitative
Mosavi F (33)	Sweden 2012	49	49/0	67	Prospective	Yes	Prostate cancer	1.5	0, 1000	DWI only	Imaging and follow-up	Qualitative
Lecouvet (19)	Belgium 2012	100	100/0	69	Prospective	Yes	Prostate cancer	1.5	ND	DWI only	Clinical judgment	Qualitative
Costelloe (20)	US 2012	29	ND	55	Prospective	ND	Breast cancer	1.5	ND	DWI with other sequence	Histology or imaging studies	Qualitative
Liang (42)	China 2010	35	11–24	56	Prospective	ND	Multi	1.5	600	DWI with other sequence	Imaging	Qualitative
Laurent V (24)	France 2010	35	35	ND	Prospective	Yes	Melanoma	1.5	0, 600	DWI only	Histology, imaging and follow-up	Qualitative
Gutzeit (21)	Switzerland 2010	36	11–25	54	Prospective	ND	Multi	1.5	1000	DWI with other sequence	Imaging	Qualitative
Zhang (43)	China 2009	18	8–10	56.2	ND	ND	Multi	3.0	800	DWI only	Imaging and follow-up	Qualitative
Wang (44)	China 2009	49	49/0	ND	Retrospective	No	Prostate cancer	3.0	800	DWI only	Imaging	Qualitative
Takenaka (34)	Japan 2009	137	83–54	72	Prospective	Yes	Lung cancer	1.5	0, 1000	DWI only + DWI with other sequence	Imaging and follow-up	Qualitative
Stecco (25)	Italy 2009	22	22 ND	ND	Prospective	Yes	Multi	1.5	1000	DWI only	Imaging	Qualitative
Xu (36)	China 2008	45	25/20	52.7	Prospective	Yes	Multi	1.5	800	DWI only	Histology, imaging and follow-up	Qualitative
Nemeth (45)	US 2007	75	75 ND	ND	ND	ND	Multi	1.5	0, 1000	DWI only	Imaging	Qualitative
Barcelo (22)	Spain 2007	24	8–16	65	Prospective	ND	Multi	1.5	600	DWI with other sequence	Imaging	Quantitative
Bhugaloo (46)	Malaysia 2006	35	13–22	62.7	Prospective	Yes	Multi	1.5	165	DWI only	Imaging and follow-up	Qualitative
Zhang (48)	China 2008	62	62 ND	69.23	ND	ND	Multi	3.0	800	DWI only	Imaging	Qualitative
Li (35)	China 2009	46	34–12	55	ND	Yes	Multi	1.5	0, 600	DWI only	Imaging	Qualitative
Pekcevik Y (41)	Turkey 2013	26	15–11	34.5	Prospective	ND	Primary bone	1.5	0, 1000	DWI with other sequence	Histology	Quantitative
Jambor I (53)	Finland 2015	53	26–27	ND	Prospective	ND	Multi	1.5	0, 150, 1000	DWI with other sequence	Imaging and follow-up	Qualitative
Ahlawat S (51)	US 2015	31	15–16	46	Retrospective	Yes	Multi	3	50, 400, 800	DWI with other sequence	Histology or imaging studies	Quantitative
Abo Dewan K.A.W (52)	Egypt 2015	50	31–19	58	ND	ND	Multi	1.5	1000	DWI with other sequence	Histology	Quantitative
Martel Villagran J (49)	Spain 2015	85	25–60	67	Prospective	ND	Multi	1.5	0, 400	DWI with other sequence	Imaging and follow-up	Qualitative
Shi (54)	China 2014	29	19–10	34	Prospective	ND	Primary bone	3	0, 300, 800	DWI only	Histology	Quantitative
Kitajima (50)	Japan 2015	62	0–62	57	Retrospective	ND	Multi	1.5	0, 1000	DWI with other sequence	Histology, clinical data	Qualitative

multi = multiple primary tumor, ND = not documented.

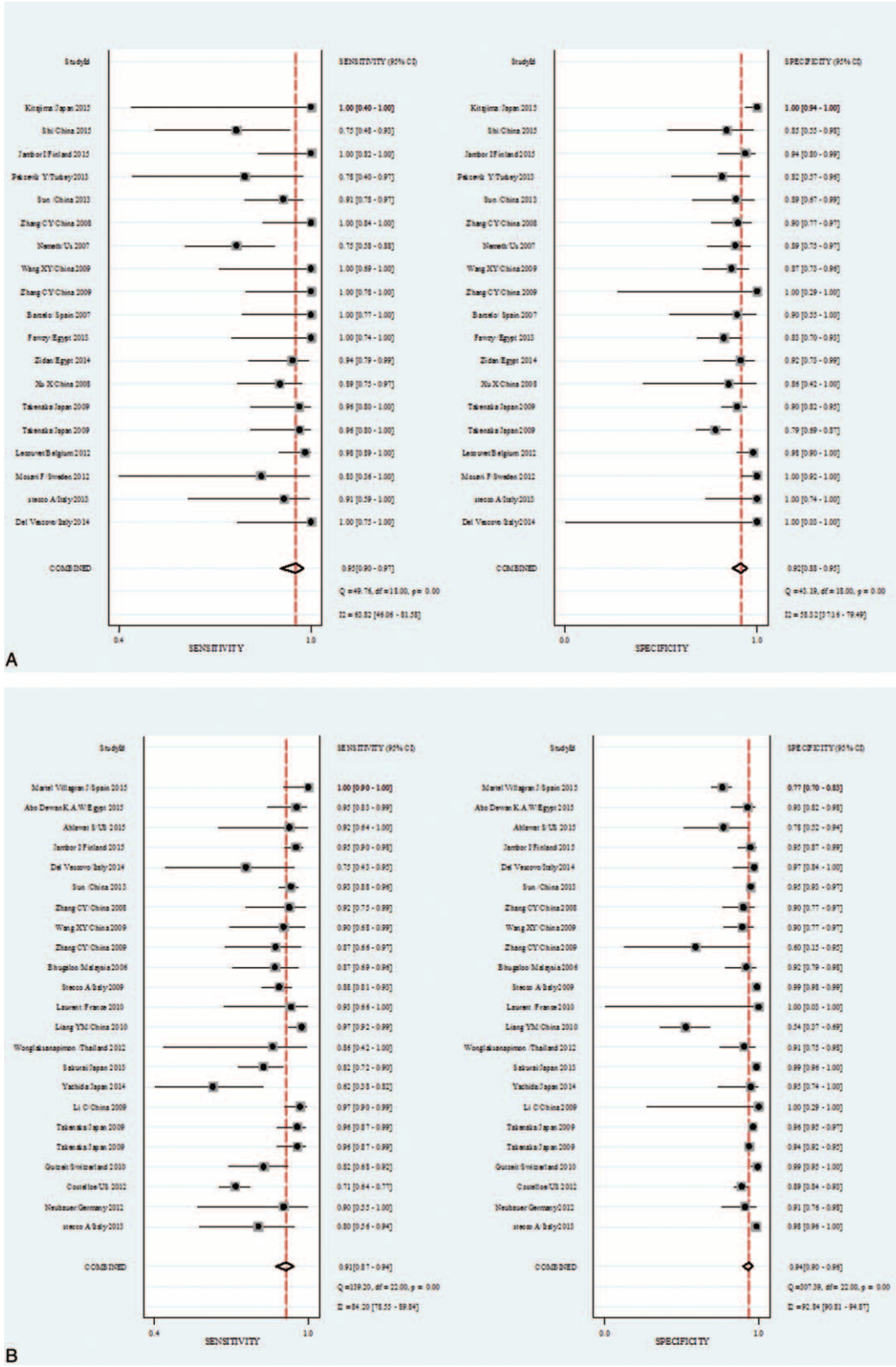

FIGURE 2. Forest plot of sensitivities and specificities on per-patient basis and per-lesion basis for the diagnosis of bone malignancy. A, Sensitivity and specificity for per-patient basis. B, Sensitivity and specificity for per-lesion basis. Each solid square represents an eligible study (error bars represent 95% CI).

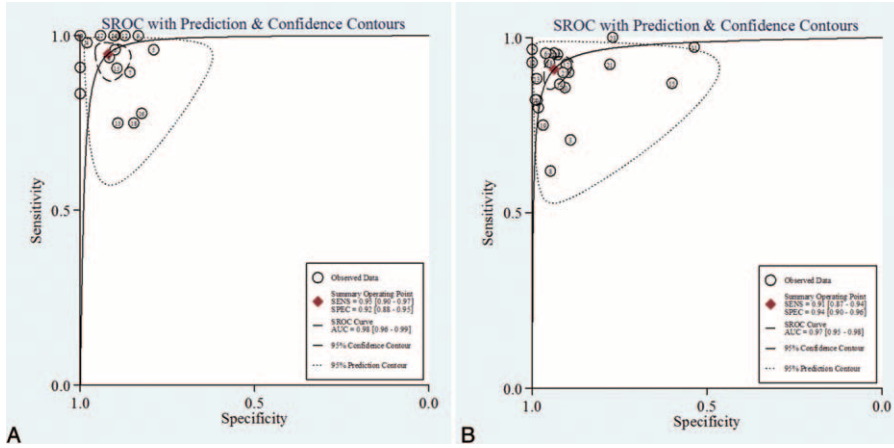

FIGURE 3. SROC curves and area under the curve of per-patient basis (A) and per-lesion basis (B) in the diagnosis of bone malignancy.

Quality Assessment of Published Studies

The quality of the included studies was assessed by the QUADAS-2 tool (Fig. 1B). Discriminations were primarily found in the domain of ''Reference Standard'' and ''Flow and Timing'' for all studies. Consequently, we selected these domains (reference standard and operation interval) as covariate in meta-regression and performed separate analyses on the subset of studies.

Overall Sensitivity, Specificity, LRs, and SROC Curves

For the assessment of efficacy of DWI in bone neoplasms (with a 95% CI reported in the included individual studies), the detailed sensitivity and specificity values on a per-lesion and per-patient basis are illustrated by the forest plot as shown in Figure 2A and B. On a per-patient basis, the pooled sensitivity, specificity were 0.95 (95% CI, 0.90–0.97) and 0.92 (95% CI, 0.88–0.95) respectively. Summary estimates indicated DOR value was 207 (95% CI, 82–523), LR^+ was 11.8 (95% CI, 7.5, 18.6), and LR^- was 0.06 (95% CI, 0.03, 0.11). The SROC curve was symmetric, and the AUC value was 0.98 (95% CI, 0.96–0.99) (Fig. 3A). These results showed that DWI provides excellent diagnostic accuracy in differentiating bone lesions on a per-patient basis.

On a per-lesion basis, the pooled sensitivity and specificity were 0.91 (95% CI, 0.87–0.94) and 0.94 (95% CI, 0.90–0.96). Summary estimates showed that DOR value was 149 (95% CI, 88–251), LR^+ was 14. 4 (95% CI, 9.1, 26.6), and LR^- was 0.10 (95% CI, 0.07, 0.14). The SROC curves were symmetric and the AUC value was 0.97 (95% CI, 0.96–0.99) (Fig. 3B). According to evaluation mentioned above, eDWI has excellent ability to both confirm and exclude presence of bone malignancy on a per-lesion basis.

The ROC space did not illustrate a curvilinear trend of points and Spearman's correlation coefficient was −0.04 ($P = 0.87$) on per-patient and 0.4 ($P = 0.06$) for a per-lesion basis. It was suggested that there was no presence of a threshold effect.

Heterogeneity and Publication Bias

eHeterogeneity was observed in sensitivity ($I^2 = 63.8$ $P < 0.01$) and specificity ($I^2 = 58.3$ $P < 0.01$) on a per-patient basis; also in sensitivity ($I^2 = 84.2$ $P < 0.01$) and specificity ($I^2 = 92.8$ $P < 0.01$) on a per-lesion basis. This result was validated by the I^2 and Cochran Q tests. Hence, the diagnostic indices were calculated using a random effect model. The results of multivariate meta-regression analysis (Table 2) demonstrated that there was significant heterogeneity in covariate of *b* value on per-lesion basis, while no significant heterogeneity in any covariate on per-patient basis.

I^2 tests also enabled us to detect heterogeneity caused by covariate of *b* value ($I^2 = 60.0$) on a per-patient basis, while covariates of standard reference ($I^2 = 52.3$) and *b* value ($I^2 = 76.7$) were responsible for the heterogeneity on a per-lesion based analysis.

The Deeks' funnel plots were generated to assess the evidence of bias toward studies (Fig. 4). According to the plot, there was no conclusive evidence of publication bias on per-patient basis ($P = 0.30$) and per-lesion basis ($P = 0.24$).

TABLE 2. Results of the Multivariable Meta-Regression Model for the Characteristics With Backward Regression Analysis (Inverse Variance Weights; Variables Were Retained in the Regression Model if $P < 0.05$)

Covariate	Coefficient	Std. Err	*P* Value	RDOR
Per-patient basis				
Consecutive	−0.557	0.8981	0.5508	0.57
Sequence	−0.120	0.7312	0.8735	0.89
Region	0.557	0.9739	0.5810	1.75
Histopathology	−0.068	0.8793	0.9402	0.93
Interval	0.896	0.7923	0.2871	2.45
Reference	−1.942	1.4869	0.2240	0.14
Quantitative	−0.462	1.9220	0.8153	0.63
Prospective	−0.254	1.1979	0.8369	0.78
b value	−2.156	1.2950	0.1289	0.11
Per-lesion basis				
Consecutive	0.873	0.4742	0.0884	2.40
Sequence	0.775	0.4793	0.1299	2.17
Region	0.384	0.5742	0.5151	1.47
Histopathology	−1.082	0.7949	0.1967	0.34
Interval	−0.495	0.3761	0.2113	0.61
Reference	0.302	0.5221	0.5728	1.35
Quantitative	−0.314	0.7730	0.6914	0.73
Prospective	0.538	0.5219	0.3213	1.71
b value	1.344	0.4600	0.0119	3.84

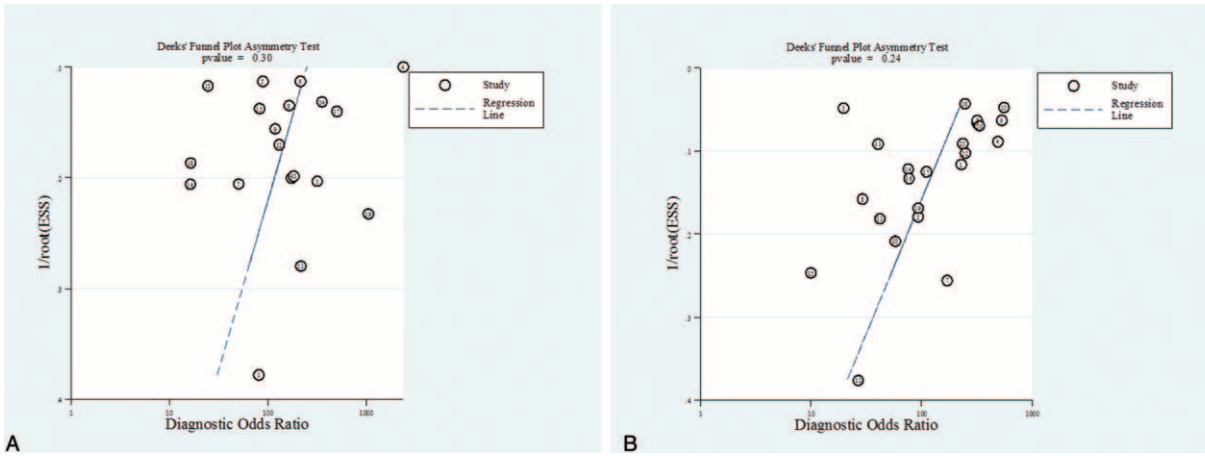

FIGURE 4. Linear regression test of funnel plot asymmetry on per-patient basis (A) and per-lesion basis (B). Each solid circle represents a study in this meta-analysis. The statistically nonsignificant *P* values of 0.30 (A) and 0.30 (B) for the slope coefficient suggest symmetry in the data.

Subgroup Analysis

We performed subgroup analysis to estimate the level of the effect by classifying studies in each covariate. The values of average adjusted sensitivity, specificity, LR^+, LR^-, DOR, and AUC of SROC curve were calculated from meta-mathematical models, which are shown in Tables 3 and 4.

According to Table 3, the results showed that the DOR and AUC in studies by qualitative analysis was significantly higher than studies that underwent quantitative analysis (combined with ADC value) on per-patient and per-lesion basis (both are $P < 0.01$).

DWI is now regarded as an adjunct to conventional MRI protocol, but would other functional imaging sequences help to further improve accuracy? The average adjusted sensitivity and specificity for MR with DWI only was 0.94 (95% CI, 0.89–0.97) and 0.89 (95% CI, 0.85–0.92), while those for DWI combined with other sequences were 0.91 (95% CI, 0.87–0.95) and 0.91 (95% CI, 0.88–0.94), the average adjusted AUC were 0.9701 and 0.9545 for DWI only and DWI combined with other sequences on a per-patient basis. When it comes to per-lesion basis, average adjusted sensitivity and specificity for DWI only were 0.90 (95% CI, 0.87–0.93) and 0.96 (95% CI, 0.95–0.97), while those for DWI combined with other sequences were 0.87 (95% CI, 0.85–0.90) and 0.94 (95% CI, 0.93–0.95), the average adjusted AUC were 0.9613 and 0.9719 for DWI only and DWI with other sequences on a per-lesion basis.

According to previous studies, we summarized *b* values that may be used as a guide when performing DWI for qualitative assessment and a *b* value in the range of 750 to 1000 s/mm^2 may be optimal for whole-body DWI. Thus, we conducted a subgroup analysis to compare the *b* value of 750 to 1000 s/mm^2 with other *b* values, pooled sensitivity, specificity, and AUC for *b* value of 750 to 1000 s/mm^2 were 0.91 (95% CI, 0.87–0.94), 0.90 (95% CI, 0.87–0.92), and 0.9535 on a per-patient basis, while those for other *b* values were 0.99 (95% CI, 0.94–1.00), 0.91 (95% CI, 0.84–0.96), and 0.9968, respectively. On a per-lesion basis, pooled sensitivity, specificity, and AUC of *b* value for 750 to 1000 s/mm^2 were 0.90 (95% CI, 0.88–0.92), 0.96 (95% CI, 0.95–0.97), and 0.9770, while those for other values were 0.85 (0.81–0.88), 0.83 (0.80–0.86), and 0.9279.

DISCUSSION

This updated study, to our knowledge, involves the largest number of patients and most comprehensive subgroup analysis in the field of whole-body DWI in differentiating bone tumors. Our findings suggest that radiation-free DWI showed a good diagnostic value on both per-patient and per-lesion-based analysis; hence, it may function as an alternative to the current nuclear medicine approach in differentiating benign from malignant bone tumors.

Despite the wide use of Tc-99m MDP BS, its use as an independent method for bony lesions is far from ideal due to lack of accuracy.[6] Another currently practiced method PET/CT also has its limitation in the detection of bone-marrow disease, as the high cellularity of normal bone marrow can be misdiagnosed as diffuse tumor infiltration or mask tumor deposits.[55] Hence, a radiation-free imaging modality with validated diagnostic accuracy is much needed for diagnosis of bone tumor.

Two previous studies indicated that the performance of DWI is similar to PET/CT, both being significantly accurate than BS in detection of bone lesion on both per-patient and per-lesion basis.[56,57] A recently published meta-analysis by Li et al showed that whole-body DWI featured similar level of sensitivity (0.897 vs. 0.895) and specificity (0.954 vs. 0.957) to PET/CT for osseous lesions detection.[56] In their study, diagnostic accuracy of DWI was compared with PET/CT in various kinds of primary and metastatic malignancies, but only a few cases were available in bone lesions. Limited sample size may impair the validity and accuracy; furthermore, only studies published in English were included which might also induce the ''Tower of Babel'' bias of their results. An earlier meta-analysis, which was published in 2011, included 11 studies with 495 patients, and the results indicated that whole-body DWI had a pooled sensitivity of 0.899 and a pooled specificity of 0.918.[57] Our updated study, by contrast, included a larger number of samples and reached similar conclusion to theirs.

The accuracy of PET/CT and scintigraphy in bone tumors differentiation can be learned from recent publications. A meta-analysis performed by Shen GH et al, including 12 published studies, provided an overview of the literature on the value of BS and PET/CT for monitoring the bone lesions.[58] The pooled sensitivity, specificity, and AUC for BS were, respectively,

TABLE 3. Quantitative Subgroup Analysis of All Available Covariate on a Per-Patient Basis

Study Characteristics		No. of Studies	I^2	Sensitivity	Specificity	Independent Estimates (95% CI)			
						LR^+	LR^-	DOR	AUC
Total		19	N/A	0.95 (0.90–0.97),	0.92 (0.88–0.95)	11.81 (7.52, 18.63)	0.06 (0.03, 0.11)	207.4 (82.4–523.0)	0.9810
Sequence of MRI	DWI with other sequence	11	0	0.91 (0.87–0.95)	0.91 (0.88–0.94)	7.80 (5.67–10.73)	0.12 (0.06–0.21)	76.9 (38.0–155.5)	0.9545
	DWI only	8		0.94 (0.89–0.97)	0.89 (0.85–0.92)	8.01 (4.29–14.94)	0.09 (0.04–0.22)	125.5 (40.6–388.6)	0.9701
Analytical method	Quantitative	4	0	0.88 (0.76–0.96)	0.84 (0.75–0.91)	5.35 (3.34–8.58)	0.18 (0.07–0.50)	30.3 (9.5–97.1)	0.9061
	Qualitative	15		0.93 (0.90–0.96)	0.91 (0.88–0.93)	8.99 (6.07–13.31)	0.09 (0.05–0.16)	115.8 (61.7–217.4)	0.9706
Study population	Caucasian	9	49	0.93 (0.88–0.96)	0.93 (0.89–0.96)	9.68 (5.53–16.96)	0.09 (0.04–0.21)	151.0 (52.1–438.0)	0.9742
	Mongoloid	10		0.92 (0.88–0.96)	0.88 (0.84–0.91)	6.54 (4.79–8.93)	0.12 (0.07–0.21)	71.1 (33.9–148.9)	0.9464
Standard reference	Reference include histopathology	6	0	0.89 (0.82–0.94)	0.90 (0.85–0.94)	6.39 (3.81–10.70)	0.16 (0.09–0.29)	53.1 (18.0–156.7)	0.9432
	Not include histopathology	13		0.94 (0.90–0.96)	0.90 (0.87–0.93)	8.30 (5.65–12.21)	0.08 (0.04–0.16)	118.4 (58.7–238.8)	0.9698
Operation interval	Less than 4 weeks	8	45	0.95 (0.90–0.98)	0.87 (0.82–0.91)	6.52 (4.62–9.19)	0.08 (0.04–0.20)	93.9 (37.2–237.2)	0.9499
	More than 4 weeks or unclear	11		0.91 (0.87–0.94)	0.93 (0.90–0.95)	8.59 (5.47–13.48)	0.12 (0.06–0.21)	105.2 (43.0–257.6)	0.9657
Same reference	Receive same reference	10	0	0.90 (0.84–0.94)	0.88 (0.83–0.92)	6.45 (4.56–9.11)	0.15 (0.08–0.28)	44.3 (20.9–93.9)	0.9367
	Receive different reference	9		0.94 (0.90–0.97)	0.91 (0.88–0.94)	10.69 (5.73–19.92)	0.09 (0.05–0.14)	175.1 (79.1–387.4)	0.9777
Consecutive enrollment	Consecutive	6	0	0.95 (0.90–0.98)	0.88 (0.85–0.92)	8.48 (4.19–17.15)	0.08 (0.04–0.18)	177.4 (64.02–491.3)	0.9795
	Inconsecutive or unclear	13		0.91 (0.87–0.94)	0.92 (0.88–0.95)	7.89 (5.58–11.15)	0.13 (0.08–0.22)	66.8 (33.4–133.5)	0.9500
Study design	Prospective	12	0	0.93 (0.90–0.96)	0.89 (0.86–0.92)	7.56 (5.04–11.34)	0.10 (0.06–0.18)	102.1 (47.0–221.5)	0.9626
	Retrospective	7		0.91 (0.84–0.96)	0.93 (0.89–0.96)	8.22 (5.17–13.09)	0.11 (0.04–0.27)	71.6 (26.5 193.9)	0.9638
b value (sec/mm^2)	750–1000	15	68	0.91 (0.87–0.94)	0.90 (0.87–0.92)	7.42 (5.49–10.03)	0.13 (0.08–0.21)	73.3 (40.0–134.3)	0.9535
	Other value or ND	4		0.99 (0.94–1.00)	0.91 (0.84–0.96)	8.81 (2.43–31.91)	0.03 (0.01–0.11)	332.5 (67.21–1644)	0.9968

Data were present as accuracy data with 95% confidence intervals. AUC = area under curve, DOR = diagnostic odds ratio, LR^- = negative likelihood ratio, LR^+ = positive likelihood ratio.

TABLE 4. Quantitative Subgroup Analysis of All Available Covariate on a Per-Lesion Basis

Study Characteristics		No. of Studies	I^2	Sensitivity	Specificity	Independent Estimates (95% CI)			
						LR^+	LR^-	DOR	AUC
Total		23	N/A	0.91 (0.87–0.94)	0.94 (0.90–0.96)	14. 4 (9.1, 22.6)	0.10 (0.07, 0.14)	148.0 (88.7–251.3)	0.9713
Sequence of MRI	DWI with other sequence	12	0	0.87 (0.85–0.90)	0.94 (0.93–0.95)	13.35 (6.88–25.90)	0.11 (0.06–0.20)	157.0 (64.9–380.1)	0.9719
	DWI only	11		0.90 (0.87–0.93)	0.96 (0.95–0.97)	10.90 (5.95–20.00)	0.12 (0.07–0.21)	101.1 (43.2–236.6)	0.9613
Analytical method	Quantitative	5	0	0.86 (0.77–0.92)	0.91 (0.85–0.95)	8.21 (5.07–13.28)	0.14 (0.04–0.47)	81.2 (30.8–213.9)	0.9593
	Qualitative	18		0.88 (0.87–0.90)	0.95 (0.94–0.96)	12.99 (7.62–22.15)	0.11 (0.07–0.17)	142.1 (69.2– 292.0)	0.9695
Study population	Caucasian	11	0	0.85 (0.82–0.87)	0.95 (0.94–0.96)	14.23 (6.45–31.40)	0.13 (0.07–0.23)	142.4 (46.36–437.7)	0.9687
	Mongoloid	12		0.92 (0.90–0.94)	0.94 (0.93–0.95)	10.38 (5.52–19.50)	0.10 (0.06–0.18)	124.9 (66.5–234.9)	0.9691
Standard reference	Reference include histopathology	6	52	0.79 (0.74–0.84)	0.85 (0.81–0.88)	6.01 (3.93–9.18)	0.10 (0.03–0.34)	61.5 (19.0–199.6)	0.9365
	Not include histopathology	17		0.91 (0.89–0.93)	0.96 (0.95–0.97)	15.11 (8.57–26.65)	0.12 (0.08–0.17)	171.2 (102.9–284.9)	0.9734
Operation interval	Less than 4 weeks	10	0	0.88 (0.85–0.90)	0.94 (0.93–0.95)	10.15 (4.93–20.88)	0.10 (0.05–0.22)	104.2 (37.7–288.4)	0.9642
	More than 4 weeks or unclear	13		0.89 (0.86–0.91)	0.96 (0.95–0.97)	13.72 (6.58–28.64)	0.13 (0.09–0.20)	186.5 (108.2–321.3)	0.9737
Same reference	Receive same reference	11	0	0.92 (0.90–0.94)	0.97 (0.6–0.98)	13.81 (4.79–39.81)	0.11 (0.08– 0.16)	150.4 (69.26–26.4)	0.9734
	Receive different reference	12		0.85 (0.82–0.88)	0.94 (0.93–0.95)	10.99 (6.75–17.90)	0.12 (0.06–0.23)	111.4 (45.14–274.8)	0.9677
Consecutive enrollment	Consecutive	9	0	0.91 (0.88–0.94)	0.96 (0.95–0.97)	14.86 (8.44–26.17)	0.11 (0.05–0.20)	174.5 (77.0–295.4)	0.9736
	Inconsecutive or unclear	14		0.87 (0.85–0.89)	0.92 (0.91–0.93)	11.14 (6.14–20.18)	0.12 (0.08–0.21)	113.9 (52.2–248.4)	0.9641
Study design	Prospective	13	0	0.88 (0.86–0.90)	0.95 (0.94–0.96)	13.71 (7.39–25.45)	0.10 (0.06–0.18)	165.2 (68.74–396.9)	0.9731
	Retrospective	10		0.89 (0.85–0.93)	0.94 (0.92–0.96)	9.60 (5.43–16.97)	0.14 (0.08–0.25)	90.2 (48.0–170.1)	0.9570
b value (s/mm^2)	750–1000	15	76	0.90 (0.88–0.92)	0.96 (0.95–0.97)	16.62 (11.11–24.87)	0.11 (0.08–0.17)	211.5 (129.6–345.2)	0.9770
	Other value or ND	8		0.85 (0.81–0.88)	0.83 (0.80–0.86)	5.41 (3.20–9.13)	0.12 (0.06–0.26)	38.1 (20.4–71.0)	0.9279

Data were present as accuracy data with 95% confidence intervals. AUC = area under curve, DOR = diagnostic odds ratio, LR^- = negative likelihood ratio, LR^+ = positive likelihood ratio.

0.79, 0.82, and 0.89 on a per-patient analysis, and respectively 0.59, 0.75, and 0.77 on a per-lesion analysis. The results of sensitivity, specificity, and AUC for PET/CT were 0.87, 0.97, and 0.95 on a per-patient analysis, while 0.83, 0.95, and 0.9494 on a per-lesion analysis. More recently, another meta-analysis conducted by Shen CT et al, including 20 articles and 1170 patients, indicated that pooled sensitivity, specificity, and AUC for BS were 0.88, 0.80, and 0.90 on a per-patient basis analysis (accuracy on a per-lesion level was not documented). Meanwhile, those for PET/CT were 0.92, 0.93, and 0.98 on a per-patient basis, and 0.87, 0.95, and 0.98 on a per-lesion analysis.[59] Both of the studies indicated a considerably lower sensitivity, specificity, and AUC of BS when compared with DWI according to our finding ($P < 0.01$). Although the specificity of DWI may be inferior to that of PET/CT, DWI featured a more superior sensitivity to PET/CT in screening bone tumor (both $P < 0.01$), and current state-of-the-art nuclear medicine imaging methods may feature insufficient sensitivity for bone tumor in general. Our results indicated that DWI may be considered a potential alternative to BS or PET/CT in screening suspicious bone malignancy for its superior sensitivity.

A factor that may affect the diagnostic accuracy of DWI and hence deriving away from the current conclusion is the different imaging protocol setup and equipment used in the different studies surveyed. Hence, we attempted to take into consideration of the varying DWI imaging conditions and to assess the level of subsequent impacts by classifying studies in each covariate with the aim of refining the validity of our research.

We assessed whether DWI combined other functional imaging sequences showed higher diagnostic accuracy than DWI alone in our stratified analysis. After adjusting for different subgroups, results showed that both of types of methods showed a similar high DOR value. Hence, DWI could function as an independent method in detecting bone malignancy.

At the present, only very limited number of research had demonstrated the predictive value of quantitative analysis with ADC value. Padhani et al found that ADC values in osseous metastasis at a cutoff greater than 0.77×10^{-3} mm^2/s resulted in a sensitivity of 0.85 and specificity of 0.90 for bone metastases differentiating from benign lesions, which was similar to our finding in quantitative analysis subgroup.[60] Compared with osseous metastasis, primary malignant bone tumors are rare and traditionally best assessed with conventional radiography for initial characterization and determination of the location of a biopsy. According to the study performed by Hayashida et al, functional imaging with ADC map alone may not be helpful for differentiating malignant tumors from benign lesions in the diagnosis of primary bone tumors.[61] According to the result of subgroup analysis, we observed that quantitative approach (based on ADC value) was slightly inferior to qualitative approach in diagnosing bone tumors. However, we noted that there were a relatively small number of quantitative approach-based studies compared with the qualitative approach-based counterparts, which may render the value of quantitative approach as assessed in this study.

In addition, DWI should be performed with appropriate choices of *b* values taking considerations of factors including anatomic region, tissue composition, and pathologic processes. Our meta-analysis showed that the *b* value in the range of 750 to 1000 s/mm^2 features higher accuracy, which is in agreement with consensus on International Society for Magnetic Resonance in Medicine Meeting.[62]

Although most of the studies included in this meta-analysis study were of good quality on the basis of QUADAS-2 criteria, heterogeneity was observed in the sensitivity and specificity in our study. A detailed meta-regression and subgroup analysis was performed to identify the potential source across studies. We expect that identification of the factors that led to the heterogeneity may strengthen the validation of our results, and the factors that influence diagnostic accuracy may help to optimize the design of future research.

We reviewed the diagnostic accuracy of DWI according to the updated methods for diagnostic meta-analyses. Despite our findings supporting high accuracy of DWI, limitations of the study affect the current strength of the evidence due to various sources of the data involved.

First, heterogeneous results of our study may affect the reliability of the conclusions. According to meta-regression analysis, although covariates of *b* value had been found as the source of the heterogeneity, we could not specify the source on a per-patient basis. Second, we compared diagnostic accuracy of DWI in our research to those of BS and PET/CT in other meta-analysis; it remains relatively inconclusive that DWI is superior to BS or PET/CT in diagnosing bone tumors. Third, most of the studies could not apply the histopathological diagnosis of bone lesions for ethical reasons because biopsies of suspected bone lesions are not part of routine examination. Therefore, part of studies used multiple imaging modalities and/or follow-up as the standards of reference instead of histopathology. Other potential limitations may be attributed to the optimization of parameters of the imaging modalities that are lack of consensus in our study.

In conclusion, our results potentially support the use of DWI as an effective method for distinguishing malignant from benign bone lesions, and utilization of radiation-free DWI may dramatically benefit population that is vulnerable to ionizing radiation. Ability to provide morphological and functional information in a single scan makes DWI attractive and promising in the diagnosis of bone tumor. Additional effort is needed for the imaging protocol standardization of DWI to achieve quality assurance.

REFERENCES

1. Parker C, Nilsson S, Heinrich D, et al. Alpha emitter radium-223 and survival in metastatic prostate cancer. *N Engl J Med.* 2013;369:213–223.
2. Chow E, van der Linden YM, Roos D, et al. Single versus multiple fractions of repeat radiation for painful bone metastases: a randomised, controlled, non-inferiority trial. *Lancet Oncol.* 2014;15:164–171.
3. Sartor O, Coleman R, Nilsson S, et al. Effect of radium-223 dichloride on symptomatic skeletal events in patients with castration-resistant prostate cancer and bone metastases: results from a phase 3, double-blind, randomised trial. *Lancet Oncol.* 2014;15:738–746.
4. Wilkinson AN, Viola R, Brundage MD. Managing skeletal related events resulting from bone metastases. *BMJ.* 2008;337:a2041.
5. Scher HI, Halabi S, Tannock I, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. *J Clin Oncol.* 2008;26:1148–1159.
6. Wong KK, Piert M. Dynamic bone imaging with 99mTc-labeled diphosphonates and 18F-NaF: mechanisms and applications. *J Nucl Med.* 2013;54:590–599.

7. Fuster D, Duch J, Paredes P, et al. Preoperative staging of large primary breast cancer with [18F]fluorodeoxyglucose positron emission tomography/computed tomography compared with conventional imaging procedures. *J Clin Oncol.* 2008;26:4746–4751.
8. Wade AA, Scott JA, Kuter I, et al. Flare response in 18F-fluoride ion PET bone scanning. *AJR Am J Roentgenol.* 2006;186:1783–1786.
9. Lin NU, Thomssen C, Cardoso F, et al. International guidelines for management of metastatic breast cancer (MBC) from the European School of Oncology (ESO)-MBC Task Force: surveillance, staging, and evaluation of patients with early-stage and metastatic breast cancer. *Breast.* 2013;22:203–210.
10. Centers for Medicare and Medicaid Services: decision memo for positron emission tomography (FDG) for solid tumors (CAG-00181R4) http://www.cms.gov/medicare-coverage-database/details/nca-decisionmemo.aspx?NCAId=26. (accessed June 11, 2013).
11. Labianca R, Nordlinger B, Beretta GD, et al. Primary colon cancer: ESMO Clinical Practice Guidelines for diagnosis, adjuvant treatment and follow-up. *Ann Oncol.* 2010;21 (suppl 5):v70–77.
12. Linet MS, Slovis TL, Miller DL, et al. Cancer risks associated with external radiation from diagnostic imaging procedures. *CA Cancer J Clin.* 2012;62:75–100.
13. National Research Council. Health risks from exposure to low levels of ionizing radiation: BEIR VII phase 2. Washington DTNAP, 2006.
14. Padhani AR. Diffusion magnetic resonance imaging in cancer patient management. *Semin Radiat Oncol.* 2011;21:119–140.
15. Padhani AR, Gogbashian A. Bony metastases: assessing response to therapy with whole-body diffusion MRI. *Cancer Imaging.* 2011;11 (Spec No A):S129–145.
16. Bains LJ, Zweifel M, Thoeny HC. Therapy response with diffusion MRI: an update. *Cancer Imaging.* 2012;12:395–402.
17. Del Vescovo R, Frauenfelder G, Giurazza F, et al. Role of whole-body diffusion-weighted MRI in detecting bone metastasis. *Radiol Med.* 2014;119:758–766.
18. Stecco A, Lombardi M, Leva L, et al. Diagnostic accuracy and agreement between whole-body diffusion MRI and bone scintigraphy in detecting bone metastases. *Radiol Med.* 2013;118:465–475.
19. Lecouvet FE, El Mouedden J, Collette L, et al. Can whole-body magnetic resonance imaging with diffusion-weighted imaging replace Tc 99m bone scanning and computed tomography for single-step detection of metastases in patients with high-risk prostate cancer? *Eur Urol.* 2012;62:68–75.
20. Costelloe CM, Kundra V, Ma J, et al. Fast Dixon whole-body MRI for detecting distant cancer metastasis: a preliminary clinical study. *J Magn Reson Imaging.* 2012;35:399–408.
21. Gutzeit A, Doert A, Froehlich JM, et al. Comparison of diffusion-weighted whole body MRI and skeletal scintigraphy for the detection of bone metastases in patients with prostate or breast carcinoma. *Skeletal Radiol.* 2010;39:333–343.
22. Barcelo J, Vilanova JC, Riera E, et al. Diffusion-weighted whole-body MRI (virtual PET) in screening for osseous metastases. *Radiologia.* 2007;49:407–415.
23. Sakurai Y, Kawai H, Iwano S, et al. Supplemental value of diffusion-weighted whole-body imaging with background body signal suppression (DWIBS) technique to whole-body magnetic resonance imaging in detection of bone metastases from thyroid cancer. *J Med Imaging Radiat Oncol.* 2013;57:297–305.
24. Laurent V, Trausch G, Bruot O, et al. Comparative study of two whole-body imaging techniques in the case of melanoma metastases: advantages of multi-contrast MRI examination including a diffusion-weighted sequence in comparison with PET-CT. *Eur J Radiol.* 2010;75:376–383.
25. Stecco A, Romano G, Negru M, et al. Whole-body diffusion-weighted magnetic resonance imaging in the staging of oncological patients: comparison with positron emission tomography computed tomography (PET-CT) in a pilot study. *Radiol Med.* 2009;114:1–17.
26. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* 2009;339:b2535.
27. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155:529–536.
28. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ.* 2004;329:168–169.
29. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol.* 1991;44:763–770.
30. Simel DL, Bossuyt PM. Differences between univariate and bivariate models for summarizing diagnostic accuracy may not be large. *J Clin Epidemiol.* 2009;62:1292–1300.
31. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol.* 2005;58:882–893.
32. Neubauer H, Evangelista L, Hassold N, et al. Diffusion-weighted MRI for detection and differentiation of musculoskeletal tumorous and tumor-like lesions in pediatric patients. *World J Pediatr.* 2012;8:342–349.
33. Mosavi F, Johansson S, Sandberg DT, et al. Whole-body diffusion-weighted MRI compared with (18)F-NaF PET/CT for detection of bone metastases in patients with high-risk prostate carcinoma. *AJR Am J Roentgenol.* 2012;199:1114–1120.
34. Takenaka D, Ohno Y, Matsumoto K, et al. Detection of bone metastases in non-small cell lung cancer patients: comparison of whole-body diffusion-weighted imaging (DWI), whole-body MR imaging without and with DWI, whole-body FDG-PET/CT, and bone scintigraphy. *J Magn Reson Imaging.* 2009;30:298–308.
35. Li C, Liu ZS, Du XM, et al. Clinical value of whole-body magnetic resonance diffusion weighted imaging on detection of malignant metastases. *Chin Med Sci J.* 2009;24:112–116.
36. Xu X, Ma L, Zhang JS, et al. Feasibility of whole body diffusion weighted imaging in detecting bone metastasis on 3.0T MR scanner. *Chin Med Sci J.* 2008;23:151–157.
37. Yachida Y, Yoshida S, Takeshita H, et al. Bone abnormal signal incidentally found in pre-biopsy diffusion-weighted MRI for suspected prostate cancer: what does it reflect? *Urol Int.* 2014;93:170–175.
38. Zidan DZ, Elghazaly HA. Can unenhanced multiparametric MRI substitute gadolinium-enhanced MRI in the characterization of vertebral marrow infiltrative lesions? *Egypt J Radiol Nucl Med.* 2014;45:443–453.
39. Sun MT, Cheng JL, Zhang Y, et al. Comparison of diffusion weighted whole body imaging with background body signal suppression and SPECT in diagnosis of bone metastasis. *Chin J Med Imaging Technol.* 2013;29:1500–1504.
40. Fawzy F, Tantawy HI, Ragheb A, et al. Diagnostic value of apparent diffusion coefficient to differentiate benign from malignant vertebral bone marrow lesions. *Egypt J Radiol Nucl Med.* 2013;44:265–271.
41. Pekcevik Y, Kahya MO, Kaya A. Diffusion-weighted magnetic resonance imaging in the diagnosis of bone tumors: preliminary results. *J Clin Imaging Sci.* 2013;3:63.
42. Liang YM, Yan CX, Zhu JZ, et al. Values of using 1.5T MR scaning in bone metastasis by WB-DWI. *Chin J Cancer Prevent Treat.* 2010;17:1123–1126.

43. Zhang CY, Ren ZQ, Sun HY, et al. Application of large field diffusion-weighted imaging in the detection of bone metastases of malignant tumors: comparison with bone scintigraphy. *Chin J Med Imaging Technol.* 2009;25:1258–1261.
44. Wang XY, Zhang CY, Jiang XX. Prospective study of bone metastasis from prostate cancer: comparison between large field diffusion-weighted imaging and bone scintigraphy. *Chin J Radiol.* 2009;43:131–135.
45. Nemeth AJ, Henson JW, Mullins ME, et al. Improved detection of skull metastasis with diffusion-weighted MR imaging. *AJNR Am J Neuroradiol.* 2007;28:1088–1092.
46. Bhugaloo A, Abdullah B, Siow Y, et al. Diffusion weighted MR imaging in acute vertebral compression fractures: differentiation between malignant and benign causes. *Biomed Imaging Interv J.* 2006;2:e12.
47. Wonglaksanapimon S, Chawalparit O, Khumpunnip S, et al. Vertebral body compression fracture: discriminating benign from malignant causes by diffusion-weighted MR imaging and apparent diffusion coefficient value. *J Med Assoc Thai.* 2012;95:81–87.
48. Zhang CY, Wang XY, Jiang XX. Combination study of large field diffusion weighted imaging and dual echo T1 weighted imaging in detection of bone metastases. *J Clin Radiol.* 2008;27:1526–1530.
49. Martel Villagran J, Bueno Horcajadas A, Perez Fernandez E, et al. Accuracy of magnetic resonance imaging in differentiating between benign and malignant vertebral lesions: role of diffusion-weighted imaging, in-phase/opposed-phase imaging and apparent diffusion coefficient. *Radiologia.* 2015;57:142–149.
50. Kitajima K, Tanaka U, Ueno Y, et al. Role of diffusion weighted imaging and contrast-enhanced MRI in the evaluation of intrapelvic recurrence of gynecological malignant tumor. *PLoS One.* 2015;10:e0117411.
51. Ahlawat S, Khandheria P, Subhawong TK, et al. Differentiation of benign and malignant skeletal lesions with quantitative diffusion weighted MRI at 3T. *Eur J Radiol.* 2015;84:1091–1097.
52. Abo Dewan KAW, Salama AAE, Habashy HMS, et al. Evaluation of benign and malignant vertebral lesions with diffusion weighted magnetic resonance imaging and apparent diffusion coefficient measurements. *Egypt J Radiol Nucl Med.* 2015;46:423–433.
53. Jambor I, Kuisma A, Ramadan S, et al. Prospective evaluation of planar bone scintigraphy, SPECT, SPECT/CT, 18F-NaF PET/CT and whole body 1.5T MRI, including DWI, for the detection of bone metastases in high risk breast and prostate cancer patients: SKELETA clinical trial. *Acta Oncol.* 2015;2:1–9.
54. Shi XY, Ren CP, Cheng JL, et al. Comparison of monoexponential and biexponential models of diffusion-weighted imaging in differential diagnosis between benign and malignant lesions of primary bone tumors. *Tumor.* 2014;34:163–168.
55. van Ufford HM, Kwee TC, Beek FJ, et al. Newly diagnosed lymphoma: initial results with whole-body T1-weighted, STIR, and diffusion-weighted MRI compared with 18F-FDG PET/CT. *AJR Am J Roentgenol.* 2011;196:662–669.
56. Li B, Li Q, Nie W, et al. Diagnostic value of whole-body diffusion-weighted magnetic resonance imaging for detection of primary and metastatic malignancies: a meta-analysis. *Eur J Radiol.* 2014;83:338–344.
57. Wu LM, Gu HY, Zheng J, et al. Diagnostic value of whole-body magnetic resonance imaging for bone metastases: a systematic review and meta-analysis. *J Magn Reson Imaging.* 2011;34:128–135.
58. Shen G, Deng H, Hu S, et al. Comparison of choline-PET/CT, MRI, SPECT, and bone scintigraphy in the diagnosis of bone metastases in patients with prostate cancer: a meta-analysis. *Skeletal Radiol.* 2014;43:1503–1513.
59. Shen CT, Qiu ZL, Han TT, et al. Performance of 18F-fluoride PET or PET/CT for the detection of bone metastases: a meta-analysis. *Clin Nucl Med.* 2015;40:103–110.
60. Padhani AR, van Ree K, Collins DJ, et al. Assessing the relation between bone marrow signal intensity and apparent diffusion coefficient in diffusion-weighted MRI. *AJR Am J Roentgenol.* 2013;200:163–170.
61. Hayashida Y, Yakushiji T, Awai K, et al. Monitoring therapeutic responses of primary bone tumors by diffusion-weighted image: initial results. *Eur Radiol.* 2006;16:2637–2643.
62. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia.* 2009;11:102–125.