



# 18F-FDG PET/CT and whole-body MRI diagnostic performance in M staging for non–small cell lung cancer: a systematic review and meta-analysis

Tássia Machado Medeiros<sup>1</sup> · Stephan Altmayer<sup>1,2</sup> · Guilherme Watte<sup>1</sup> · Matheus Zanon<sup>2,3</sup> · Adriano Basso Dias<sup>2</sup> · Natália Henz Concatto<sup>4</sup> · Julia Hoefel Paes<sup>1</sup> · Rita Mattiello<sup>1</sup> · Francisco de Souza Santos<sup>1</sup> · Tan-Lucien Mohammed<sup>5</sup> · Nupur Verma<sup>5</sup> · Bruno Hochegger<sup>1,2,3</sup>

Received: 19 August 2019 / Revised: 26 December 2019 / Accepted: 3 February 2020 / Published online: 3 March 2020  
© European Society of Radiology 2020

## Abstract

**Objectives** To evaluate the diagnostic test accuracy of 18F-fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT), whole-body magnetic resonance imaging (WB-MRI), and whole-body diffusion-weighted imaging (WB-DWI) for the detection of metastases in patients with non–small cell lung cancer (NSCLC).

**Methods** MEDLINE, Embase, and Cochrane Library databases were searched up to June 2019. Studies were selected if they reported data that could be used to construct contingency tables to compare 18F-FDG PET/CT, WB-MRI, and WB-DWI. Two authors independently extracted data on study characteristics and assessed methodological quality using the Quality Assessment of Diagnostic Accuracy Studies. Forest plots were generated for sensitivity and specificity of 18F-FDG PET/CT, WB-MRI, and whole-body diffusion-weighted imaging (WB-DWI). Summary receiver operating characteristic plots were created.

**Results** The 4 studies meeting inclusion criteria had a total of 564 patients and 559 lesions, 233 of which were metastases. In studies of 18F-FDG PET/CT, the pooled estimates of sensitivity and specificity were 0.83 (95% confidence interval [CI], 0.54–0.95) and 0.93 (95% CI, 0.87–0.96), respectively. For WB-MRI, pooled sensitivity was 0.92 (95% CI, 0.18–1.00) and pooled specificity was 0.93 (95% CI, 0.85–0.95). Pooled sensitivity and specificity for WB-DWI were 0.78 (95% CI, 0.46–0.93) and 0.91 (95% CI, 0.79–0.96), respectively. There was no statistical difference between the diagnostic odds ratio of WB-MRI and WB-DWI compared with that of PET/CT ( $p = 0.186$  for WB-DWI;  $p = 0.638$  for WB-MRI).

**Conclusion** WB-MRI and DWI are radiation-free alternatives with comparable diagnostic performance to 18F-FDG PET/CT for M staging of NSCLC.

## Key Points

- Whole-body MRI with or without diffusion-weighted imaging has a high accuracy for the diagnostic evaluation of metastases in patients with non–small cell lung cancer.
- Whole-body MRI may be used as a non-invasive and radiation-free alternative to positron emission tomography with CT with similar diagnostic performance.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00330-020-06703-1>) contains supplementary material, which is available to authorized users.

✉ Bruno Hochegger  
brunoho@ufcspa.edu.br

<sup>1</sup> Postgraduate Program in Medicine and Health Sciences, Pontifícia Universidade Católica do Rio Grande do Sul, Av. Ipiranga, 6690, Porto Alegre 90619900, Brazil

<sup>2</sup> Medical Imaging Research Lab, LABIMED, Department of Radiology, Pavilhão Pereira Filho Hospital, Irmandade Santa Casa de Misericórdia de Porto Alegre, Av. Independência, 75, Porto Alegre 90020160, Brazil

<sup>3</sup> Department of Radiology, Federal University of Health Sciences of Porto Alegre, R. Sarmento Leite, 245, Porto Alegre 90050170, Brazil

<sup>4</sup> Department of Radiology, Hospital de Clínicas de Porto Alegre, R. Ramiro Barcelos, 2350, Porto Alegre 90035903, Brazil

<sup>5</sup> Department of Radiology, College of Medicine - University of Florida, 1600 SW Archer Rd, Gainesville, FL 32611, USA

**Keywords** Positron emission tomography/computed tomography · Magnetic resonance imaging · Lung cancer · MRI

### Abbreviations

18F-FDG PET/CT	18F-Fluorodeoxyglucose positron emission tomography/computed tomography
ADC	Apparent diffusion coefficient
DWI	Diffusion-weighted imaging
EQUATOR	Enhancing the Quality and Transparency of Health Research
FP	False positive
NSCLC	Non-small cell lung cancer
PRISMA	Preferred Reporting Items for Systematic Reviews
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
SUV	Standardized uptake value
TN	True positive
WB-DWI	Whole-body diffusion-weighted imaging
WB-MRI	Whole-body magnetic resonance imaging

### Introduction

Lung cancer is the leading cause of cancer-related death worldwide, with approximately 40% of patients having distant metastases at the time of initial diagnosis [1, 2]. Because appropriate staging is crucial for decisive treatment [3], an accurate and cost-effective method for lung cancer staging must be established.

Hybrid imaging using positron emission tomography and computed tomography with fluorine-18-fluorodeoxyglucose (18F-FDG PET/CT) tracer is a powerful tool for initial staging and restaging of lung cancer, as it combines metabolic and anatomic data [3]. Furthermore, 18F-FDG PET/CT can provide surgical and radiotherapy guidance and help predict tumor response to treatment [3]. Previous investigations show that hybrid 18F-FDG PET/CT is more effective than computed tomography (CT) or positron emission tomography alone for tumor, node, metastasis staging [4]. However, inherent limitations include high cost, lack of necessary infrastructure in many centers, and high rates of false positives in areas with endemic granulomatous diseases [3].

Whole-body magnetic resonance imaging (WB-MRI) is a non-invasive and radiation-free imaging tool for cancer staging and metastasis detection [5]. Inclusion of diffusion-weighted imaging sequences with WB-MRI (WB-DWI) can further improve diagnostic accuracy [6, 7]. A recent meta-analysis compared the diagnostic performance of 18F-FDG PET/CT and DWI in differentiating malignant and benign pulmonary nodules and masses [8]. Overall, DWI appeared to be equivalent or superior to 18F-FDG PET/CT in classifying malignant lung nodules [8]. Taylor et al recently reported

that WB-MRI staging has similar accuracy to current standard methods, reducing staging costs and time [9]. However, there have been no meta-analyses comparing these imaging modalities with respect to diagnostic performance in lung cancer staging and detection of distant metastases. The aim of this systematic review and meta-analysis was to compare the diagnostic performance of 18F-FDG PET/CT, WB-MRI, and WB-DWI in the detection of extrathoracic lung cancer metastases.

### Materials and methods

#### Literature search

This study was performed using the Enhancing the Quality and Transparency of Health Research (EQUATOR) Reporting Guidelines with the Preferred Reporting Items for Systematic Reviews (PRISMA). We gathered all accessible literature available through PubMed (U.S. National Library of Medicine), Embase (Elsevier), and the Cochrane Library (John Wiley & Sons) electronic databases up to June 2019. The search algorithm was based on a combination of the equivalent terms listed in Supplementary File 1.

#### Inclusion and exclusion criteria

To be included, studies had to meet several criteria: (i) performance evaluation of 18F-FDG PET/CT or WB-MRI in M staging of lung cancer; (ii) use of histopathologic analysis or imaging follow-up as the reference standard; and (iii) inclusion of clearly stated values for true positive (TN), false positive (FP), false negative (FN), and true negative (TN). Studies were excluded if they (i) focused on prognosis or therapeutic response instead of M staging; (ii) had a sample of fewer than 10 patients; (iii) were published as a conference abstract, letter, review, animal experiment, comment, or case report; (iv) were not published in English; (v) used 18F-FDG PET that was not hybridized with CT; (vi) used radiotracers other than 18F-FDG; or (vii) used non-whole body MRI. Three researchers reviewed the titles and abstracts of retrieved articles and applied inclusion and exclusion criteria. The full texts of qualifying articles were retrieved and reviewed to confirm study eligibility. The PRISMA flowchart for the selection process is presented in Fig. 1.

#### Assessment of methodologic quality

Studies that met eligibility criteria were examined by 2 reviewers following the Quality Assessment of Diagnostic

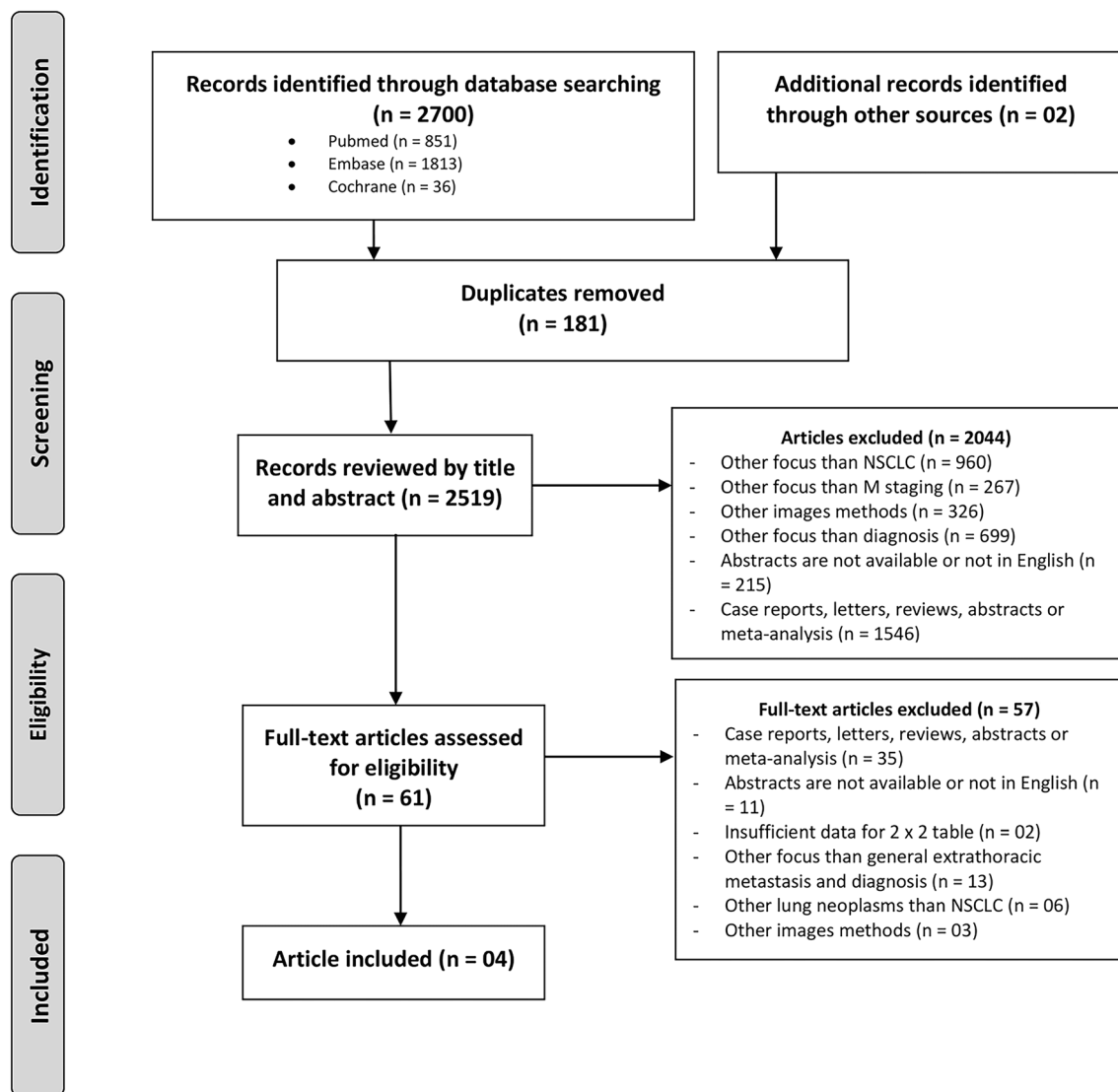


Fig. 1 Study flow diagram of literature search of eligible studies

Accuracy Studies (QUADAS-2) [10]. This quality control instrument consists of 4 parts: patient selection, index testing, reference standard, and flow and timing. The final criterion is based on the risk of bias with respect to concerns about applicability. Rating risks of bias was determined as high, low, or unclear. Only studies with a low risk of bias were included in the present study. Disagreements between reviewers were resolved by consensus.

### Data extraction

Literature accepted for analysis was reviewed by 2 analysts using the PRISMA guidelines [11]. Information collected from studies included first author, year of publication, study design, country of patient recruitment, patient enrollment, technical specifications, reference standard, and blinding.

Details regarding staging, and numbers of TP, TN, FP, and FN were also gathered from each article.

### Statistical analysis

Studies were only included if they included both WB-MRI with or without DWI and 18F-FDG PET/CT in the same sample group so as to minimize methodologic and clinical inter-study heterogeneity. Pooled sensitivities and specificities with 95% confidence intervals (95% CIs) were calculated using the bivariate random effect analysis model of Reitsma et al [12]. Pooled estimates of positive and negative likelihood ratios (PLRs and NLRs) and diagnostic odds ratios (DORs) were also obtained. Direct comparison of the DORs of MRI and PET/CT were performed using a Z test, and a two-tailed *p* value of less than 0.05 was considered significant [13–15]. Given that the DOR does not follow a Gaussian distribution,

we transformed the natural logarithm of DOR for the purpose of this analysis to assume an approximately normal distribution [15]. Summary receiver operating characteristic curves were constructed, and areas under the curve were obtained. Heterogeneity among studies was assessed using the chi-square statistic for the pooled estimates ( $p < 0.05$  indicated significant heterogeneity). We further calculated the heterogeneity of the pooled estimate of DORs to test if heterogeneity was due to the threshold effect. The variation across studies caused by heterogeneity rather than by chance was estimated by calculating the  $I^2$  values. Deeks' funnel plot was intended to assess for publication bias, as indicated by an asymmetric appearance [16]. Analyses were conducted using Stata version 12.0 (StataCorp LP).

## Results

### Literature search

The initial literature search resulted in 2700 articles, of which 61 were reviewed and 4 were considered eligible. Ohno et al [17] and Chen et al [18] compared 18F-FDG PET/CT with WB-DWI. Ohno and colleagues [14] also performed comparisons with WB-MRI, as did Ohno et al and Yi et al [19, 20]. Because of our focus on M staging and WB-MRI with or without DWI, the other methods analyzed by these authors will not be referenced, and diagnostic capability will include brain imaging.

Summary findings of the eligible studies are shown in Table 1. The four studies included a total of 553 patients, of those at least 87 had M-stage lung cancer. Chen et al did not report the exact number of patients with metastases, and they were the only authors that used the total number of metastases as the reference standard to calculate diagnostic accuracy (instead of the number of patients with M-stage NSCLC as the others) [18]. Technical characteristics of the eligible studies

(equipment, sequences, diagnostic parameters) are described in Supplementary Table 1. All WB-MRI studies were performed with the use of contrast media [17, 19, 20]. Both WB-DWI studies used  $b$  values of 0 and 1000 s/mm<sup>2</sup> [17, 18]. Ohno et al's work [17] was the only study that used a 5-point visual scale (positive if visual scale  $\geq 4$ ) as a threshold for probability of malignancy rather than consensus of two radiologists. All studies enrolled Asian patients. A summary of lesion descriptions is shown in Supplementary Table 2.

### Methodological quality

Participant selection was considered at low risk of bias in all studies. Regarding the reference standard, most studies were judged as low risk of bias because they used histopathologic analysis and a follow-up of more than 12 months [17] or more than 6 months [18–20]. Studies could not be evaluated with respect to risk of bias for flow and timing, as the time intervals between the index tests and reference standard tests were not reported. The results of the QUADAS-2 assessment are presented in Supplementary Figs. 1 and 2.

### Heterogeneity between studies and publication bias

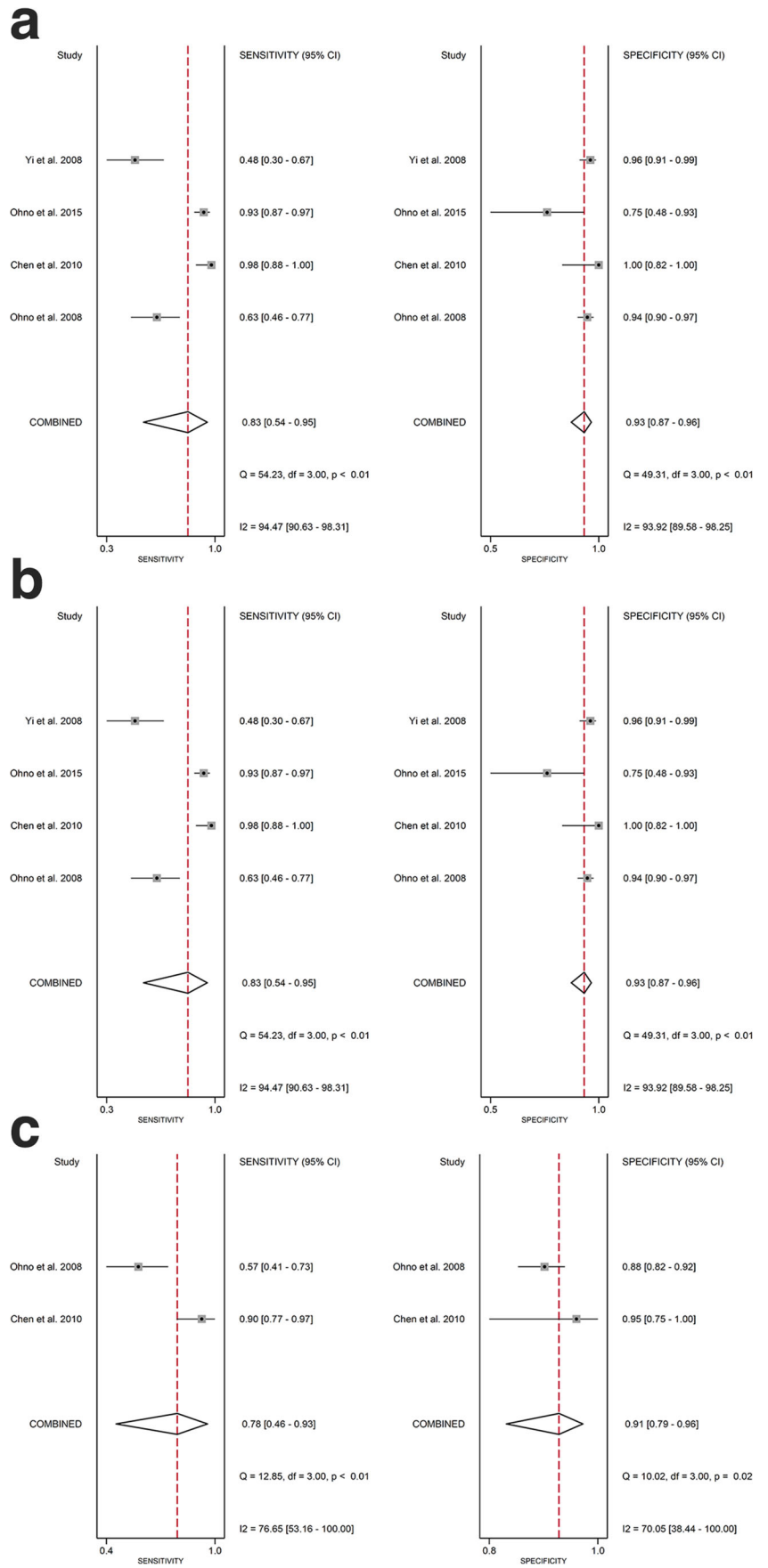
All four studies included in this meta-analysis exhibited significant heterogeneity ( $p < 0.01$ ) with respect to sensitivity and specificity for 18F-FDG PET/CT, WB-MRI, and WB-DWI. For WB-DWI, the specificity  $p$  value was less than 0.02. Specificity heterogeneity showed strong variability for 18F-FDG PET/CT and WB-MRI ( $I^2$  of 93.9% and 90.9%, respectively) and moderate variability for WB-DWI (70%). When measured for the pooled estimate of DORs, heterogeneity was not statistically significant for PET/CT ( $I^2 = 37.9%$ ,  $p = 0.185$ ), WB-MRI ( $I^2 = 57.0%$ ,  $p = 0.098$ ), or WB-DWI ( $I^2 = 0%$ ,  $p = 0.317$ ).

**Table 1** Characteristics of the studies included in the meta-analysis

	Ohno et al [17]	Chen et al [18]	Ohno et al [19]	Yi et al [20]
Year	2008	2010	2015	2008
Country	Japan	China	Japan	Korea
Study design	P	P	P	P
Patient enrollment	C	C	C	C
Mean age (range)	72 (47, 85)	51 (35, 76)	72.1 (47, 83)	61 (34, 82)
Reference standard	His/Fol	His/Fol	His/Fol	His/Fol
Blinding	Yes	Yes	Yes	Yes
Total sample/patients with metastases (number)	203/40	56/NA	140/16	154/31
Metastatic lesions ( $n$ )	93	43	NA	49

C, consecutive; Fol, follow-up; His, histological; NA, not available; P, prospective

**Fig. 2** Forest plot of the 18F-FDG PET/CT (a), WB-MRI (b), and WB-DWI (c) studies. The  $Q$  statistic and  $I^2$  are measurements of heterogeneity. Heterogeneity was strong and significant for all pooled analyses, except for specificity for WB-DWI





**Table 2** Summary of test performance characteristics based on diagnostic capability on assessment of metastases of integrated 18F-FDG PET/CT, WB-DWI, and WB-MRI without DWI included brain analysis

Parameter	WB-DWI ( <i>n</i> = 02)	95% CI	WB-MRI ( <i>n</i> = 03)	95% CI	18F-FDG PET/CT ( <i>n</i> = 04)	95% CI
TP	61	-	164	-	197	-
TN	162	-	280	-	303	-
FP	21	-	22	-	18	-
FN	21	-	31	-	41	-
Sensitivity (%)	78	46, 93	92	18, 100	83	54, 95
Specificity (%)	91	79, 96	92	85, 95	93	87, 96
PPV (%)	88	82, 95	90	86, 94	91	87, 95
NPV (%)	79	73, 84	90	86, 94	83	79, 86
PLR	8.7	2.9, 25.6	10.8	6.4, 18.4	11.7	6.6, 20.9
NLR	0.24	0.08, 0.77	0.09	0.00, 3.25	0.19	0.06, 0.58
DOR	35	4, 279	117	3, 4480	62	18, 212
AUC	0.93	0.90, 0.95	0.93	0.91, 0.95	0.95	0.93, 0.96

AUC, area under the curve; CI, confidence interval; DOR, diagnostic odds ratio; DWI, diffusion-weighted imaging; FN, false negative; FP, false positive; PLR, positive likelihood ratio; PPV, positive predictive value; NLR, negative likelihood ratio; NPV, negative predictive value; TN, true negative; TP, true positive; WB-MRI, whole-body magnetic resonance imaging; 18F-FDG PET/CT, 18F-fluorodeoxyglucose positron emission tomography/computed tomography

### Diagnostic accuracy of 18F-FDG PET/CT, WB-MRI, and WB-DWI

Pooled results are shown in Fig. 2. 18F-FDG PET/CT had a pooled sensitivity of 83% (95% CI, 0.54–0.95) and specificity of 93% (95% CI, 0.87–0.96). WB-MRI had a pooled sensitivity of 92% (95% CI, 0.18–1.00) and specificity of 92% (95% CI, 0.85–0.95), whereas WB-DWI had a pooled sensitivity of 78% (95% CI, 0.46–0.93) and specificity of 0.91 (95% CI, 0.79–0.96). The likelihood ratio syntheses resulted in an overall PLR of 8.7 (95% CI, 2.9–25.6) and NLR of 0.24 (95% CI, 0.08–0.77) for WB-DWI and a PLR of 10.8 (95% CI, 6.4–18.4) and NLR of 0.09 (95% CI, 0.00–3.25) for WB-MRI. For 18F-FDG PET/CT, the overall PLR was 11.7 (95% CI, 6.6–20.9), and the NLR was 0.19 (95% CI, 0.06–0.58). The DOR was 62 (95% CI, 18–212) for WB-DWI, 117 (3–4480) for WB-MRI, and 62 (95% CI, 18–212) for 18F-FDG PET/CT (Table 2). Direct comparison of the DORs of MRI to PET/CT revealed no statistical significance between imaging modalities ( $p = 0.186$  for WB-DWI;  $p = 0.638$  for WB-MRI). Using a fitted summary receiver operating characteristic curve, the overall areas under the curve for WB-DWI, WB-MRI, and 18F-FDG PET/CT were 0.93 (95% CI, 0.90–0.95), 0.93 (95% CI, 0.91–0.95), and 0.95 (95% CI, 0.93–0.96), respectively (Fig. 3).

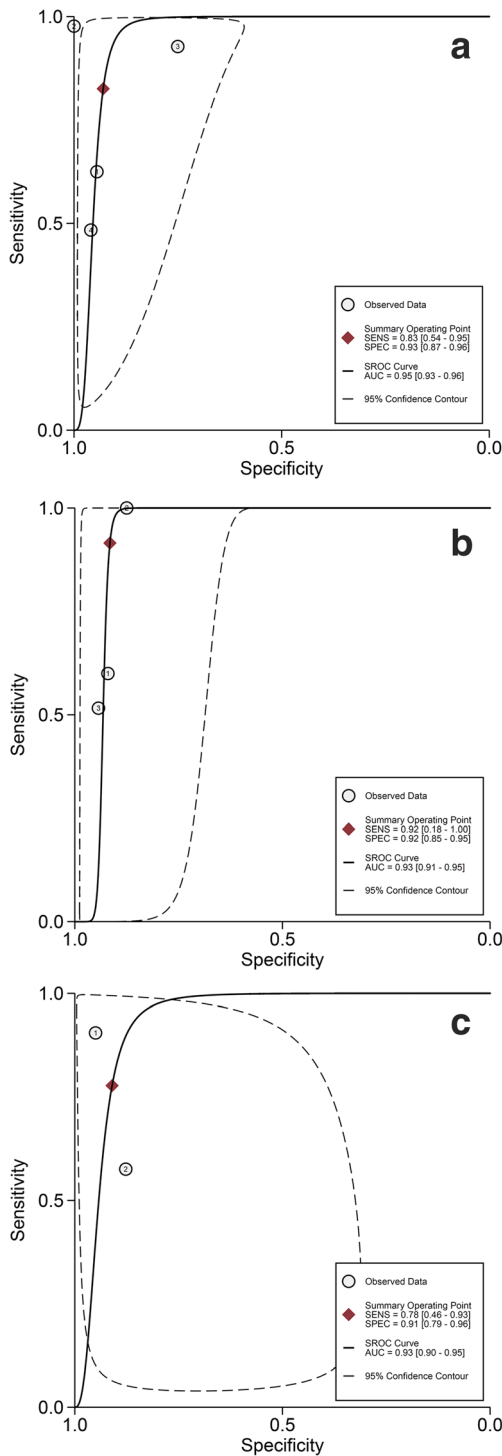
### Discussion

MRI has a high potential to be a single-test imaging modality for evaluation of NSCLC patients because of its comparable

accuracy to PET/CT, reduced examination time, and lack of ionizing radiation [9, 21, 22]. Also, MRI is more cost-effective for reduction of health care costs as it usually costs half the price of having a PET/CT study [23]. Previous meta-analyses have demonstrated that both modalities (MRI and 18F-FDG PET/CT) have a good diagnostic performance in evaluating pulmonary lesions, lymph nodes in non-small cell lung cancer (NSCLC), and detection of primary and metastatic malignancies [7–9, 24, 25]. Our results summarize those few studies that have focused on global M staging in NSCLC using both imaging techniques. Only four studies fulfilled our study criteria [17–20], and two of those used DWI [17, 18].

The heterogeneity found in the initial analysis (sensitivity, specificity) was most likely due to the threshold effect, as it did not remain significant in the analysis of the pooled estimates of DORs. Other possible explanations for heterogeneity included different sample sizes, the impact of per patient analysis instead of per lesion analysis, varying composition of organ and tumor histopathology, and endemic zones of granulomatous disease in China and Korea [26–28]. Our analysis showed that the studied techniques had similar probabilities of ruling out malignancy (NLRs) or positive results among those with disease (PLRs). For a test to be highly useful, it should have an NLR less than 0.1 and a PLR greater than 10. Thus, WB-MRI would be more highly indicated (NLR, 10.8; PLR, 0.09), and WB-DWI should not be used alone (NLR, 8.7; PLR, 0.24) [7]. The DOR, which measures discriminative power, did not differ between diagnostic tests.

Because there was no differences in overall diagnostic performance (i.e., the DORs) between WB-MRI and 18F-FDG PET/CT, WB-MRI appears to be a suitable, accurate



**Fig. 3** Summary ROC (SROC) curve of the 18F-FDG PET/CT (a), WB-MRI (b), and WB-DWI (c) studies

alternative to 18F-FDG PET/CT. The use of DWI may provide supplemental information for decision-making in M staging of NSCLC [17]. Ohno et al used 4 modalities (WB-DWI only, WB-MRI without DWI, WB-MRI plus DWI, and 18F-FDG PET/CT) to evaluate lesions. They found that, when brain metastasis assessment was included, specificity and

accuracy were lower with WB-DWI alone (87.7% and 81.8%) than with WB-MRI plus DWI (92% and 87.7%), WB-MRI without DWI (92% and 85.7%), and 18F-FDG PET/CT (94.5% and 88.2%) [17]. However, Chen et al [18] demonstrated no difference in the diagnostic performance of WB-DWI and 18F-FDG PET/CT for the detection of metastases. Yi et al [20] found no difference in detection ability for brain and hepatic metastases between WB-MRI and 18F-FDG PET/CT. In comparing the diagnostic accuracy of extrathoracic metastases, Ohno et al [19] found WB-MRI to be superior to 18F-FDG PET/CT (98.6% vs. 90.7%,  $p < 0.05$ ) but not to FDG-PET/MRI.

The incidence of brain metastases in patients with NSCLC ranges from 21 to 54% and increases as overall survival increases [29]. Because 18F-FDG PET/CT provides limited information (inferior soft tissue contrast and high physiologic background activity), brain MRI is the preferred and recommended imaging modality for patients with suspected NSCLC [30]. Ohno et al [17] assessed the actual utility of WB-MRI compared with 18F-FDG PET/CT. This study demonstrated that the diagnostic capability for M staging, excluding brain metastasis evaluation, was inferior for WB-MRI without DWI compared with 18F-FDG PET/CT. Lee et al [31] showed that 18F-FDG PET/CT plus contrast-enhanced brain MRI without DWI had a higher sensitivity than 18F-FDG PET/CT alone (88% vs. 24%;  $p < 0.001$ ) to detect brain metastases in patients with lung adenocarcinoma.

Bone is the site of 30 to 40% of lung cancer metastases [32], and bone metastasis prevalence was within this range in all 4 of the studies included in this meta-analysis. Ohno et al [17] and Chen et al [18] found that 18F-FDG PET/CT and WB-MRI had a similar performance in detecting bone metastases. Yi et al [20] did not perform DWI but did obtain additional T1-weighted turbo spin-echo images that showed bones to be most frequent site of metastasis (8%). Takenaka et al [33] showed that WB-MRI with or without DWI is more specific and accurate in detecting bone metastases compared with WB-DWI alone, 18F-FDG PET/CT, or bone scintigraphy on a per-site basis in patients with NSCLC. They also concluded that adoption of DWI as an adjunct for WB-MRI could improve the diagnostic accuracy.

Our study has some limitations. The number of articles and patients examined was smaller than anticipated. In diagnostic imaging studies, small sample sizes and heterogeneous methods of primary studies can limit the quality of the meta-analysis [34]. We were not able to test for publication bias given the small number of studies included in the meta-analysis. Other limitations include those inherent to this study design, such as selection and publication bias, limited information from reports, and potential for ecological fallacy. Description of characteristics of metastatic lesions (e.g., size, ADC, and SUV value) was not available in most studies. Last, exclusion of non-English studies may have increased the probability of publication bias.

## Conclusions

This meta-analysis of four separate studies found WB-MRI and WB-DWI to show a similar diagnostic performance to 18F-FDG PET/CT in M-staging of NSCLC. These MRI techniques are of lower cost and less time-consuming than PET/CT and are ionizing-radiation-free. Further high-quality studies comparing the diagnostic performance of these imaging modalities and various optimized MRI protocols are needed to determine if MRI should supplant the current standard approach to M staging.

**Acknowledgments** We acknowledge the contribution of the Coordenacao de Aperfeicoamento de Pessoal de Nivel Superior - Brazil (CAPES).

**Funding Information** The authors state that this work has not received any funding.

## Compliance with ethical standards

**Guarantor** The scientific guarantor of this publication is Bruno Hochhegger, MD, PhD.

**Conflict of interest** The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

**Statistics and biometry** One of the authors has significant statistical expertise.

**Informed consent** Written informed consent was not required for this study because this represents a literature review.

**Ethical approval** Institutional Review Board approval was not required because this represents a literature review and meta-analysis.

## Methodology

- Diagnostic or prognostic study
- Multicenter study

## References

- World Health Organization (2016) Fact sheets. Cancer. World Health Organization, Geneva. Available via <http://www.who.int/news-room/fact-sheets/detail/cancer>. Accessed 1 Feb 2018
- Xiang D, Zhang B, Doll D, Shen K, Kloecker G, Freter C (2013) Lung cancer screening: from imaging to biomarker. *Biomark Res* 1: 1–9. <https://doi.org/10.1186/2050-7771-1-4>
- Liam CK, Andarini S, Lee P, Ho JC, Chau NQ, Tscheikuna J (2015) Lung cancer staging now and in the future. *Respirology* 20:526–534. <https://doi.org/10.1111/resp.12489>
- Bruzzi JF, Munden RF (2006) PET/CT imaging of lung cancer. *J Thorac Imaging* 21:123–136
- Berger A (2002) Magnetic resonance imaging. *BMJ* 324:35
- Lee MH, Kim SR, Park SY et al (2012) Application of whole-body MRI to detect the recurrence of lung cancer. *Magn Reson Imaging* 30:1439–1445. <https://doi.org/10.1016/j.mri.2012.04.014>
- Li B, Li Q, Nie W, Liu S (2014) Diagnostic value of whole-body diffusion-weighted magnetic resonance imaging for detection of primary and metastatic malignancies: a meta-analysis. *Eur J Radiol* 83:338–344. <https://doi.org/10.1016/j.ejrad.2013.11.017>
- Dias AB, Zanon M, Altmayer S et al (2019) Fluorine 18-FDG PET/CT and diffusion-weighted MRI for malignant versus benign pulmonary lesions: a meta-analysis. *Radiology* 290:525–534. <https://doi.org/10.1148/radiol.2018181159>
- Taylor SA, Mallett S, Ball S et al (2019) Diagnostic accuracy of whole-body MRI versus standard imaging pathways for metastatic disease in newly diagnosed non-small-cell lung cancer: the prospective Streamline L trial. *Lancet Respir Med* 7:523–532. [https://doi.org/10.1016/S2213-2600\(19\)30090-6](https://doi.org/10.1016/S2213-2600(19)30090-6)
- Whiting PF, Rutjes AWS, Westwood ME et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 155:529–536. <https://doi.org/10.7326/0003-4819-155-8-201110180-00009>
- Moher D, Liberati A, Tetzlaff J, Altman DG (2010) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 8:336–341. <https://doi.org/10.1016/j.ijsu.2010.02.007>
- Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH (2005) Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 58:982–990. <https://doi.org/10.1016/J.JCLINEPI.2005.02.022>
- Hanley JA, McNeil BJ (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 148(3):839–843. <https://doi.org/10.1148/radiology.148.3.6878708>
- Kiewiet JJ, Leeuwenburgh MM, Bipat S, Bossuyt PM, Stoker J, Boermeester MA (2012) A systematic review and meta-analysis of diagnostic performance of imaging in acute cholecystitis. *Radiology* 264(3):708–720. <https://doi.org/10.1148/radiol.12111561>
- Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM (2003) The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 56(11):1129–1135. [https://doi.org/10.1016/s0895-4356\(03\)00177-x](https://doi.org/10.1016/s0895-4356(03)00177-x)
- Deeks JJ, Macaskill P, Irwig L (2005) The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 58:882–893. <https://doi.org/10.1016/j.jclinepi.2005.01.016>
- Ohno Y, Koyama H, Onishi Y et al (2008) Non-small cell lung cancer: whole-body MR examination for M-stage assessment-utility for whole-body diffusion-weighted imaging compared with integrated FDG PET/CT. *Radiology* 248:643–654. <https://doi.org/10.1148/radiol.2482072039>
- Chen W, Jian W, Li HT et al (2010) 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. <https://doi.org/10.1016/j.mri.2010.02.009>
- Ohno Y, Koyama H, Yoshikawa T et al (2015) Three-way comparison of whole-body MR, coregistered whole-body FDG PET/MR, and integrated whole-body FDG PET/CT imaging: TNM and stage assessment capability for non-small cell lung cancer patients. *Radiology* 275:849–861. <https://doi.org/10.1148/radiol.14140936>
- Yi CA, Shin KM, Lee KS et al (2008) Non-small cell lung cancer staging: efficacy comparison of integrated PET/CT versus 3.0-T whole-body MR imaging. *Radiology* 248:632–642. <https://doi.org/10.1148/radiol.2482071822>
- Morone M, Bali MA, Tunariy N et al (2017) Whole-body MRI: current applications in oncology. *AJR Am J Roentgenol* 209: W336–W349. <https://doi.org/10.2214/AJR.17.17984>
- Usuda K, Funazaki A, Maeda R et al (2017) Economic benefits and diagnostic quality of diffusion-weighted magnetic resonance



- imaging for primary lung cancer. *Ann Thorac Cardiovasc Surg* 23: 275–280. <https://doi.org/10.5761/atcs.ra.17-00097>
23. Plathow C, Walz M, Lichy MP et al (2008) Kostenüberlegungen zur Ganzkörper-MRT und PET-CT im Rahmen des onkologischen Stagings. *Radiologe* 48:384–396. <https://doi.org/10.1007/s00117-007-1547-z>
  24. Shen G, Lan Y, Zhang K, Ren P, Jia Z (2017) Comparison of 18F-FDG PET/CT and DWI for detection of mediastinal nodal metastasis in non-small cell lung cancer: a meta-analysis. *PLoS One* 12:1–18. <https://doi.org/10.1371/journal.pone.0173104>
  25. Wu LM, Xu JR, Gu HY et al (2012) 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. <https://doi.org/10.1016/j.jss.2012.03.074>
  26. Devillé WL, Buntinx F, Bouter LM et al (2002) Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2:9. <https://doi.org/10.1186/1471-2288-2-9>
  27. Deppen SA, Blume JD, Kensinger CD et al (2014) Accuracy of FDG-PET to diagnose lung cancer in areas with infectious lung disease: a meta-analysis. *JAMA* 312:1227–1236. <https://doi.org/10.1001/jama.2014.11488>
  28. World Health Organization (2018) Global tuberculosis report 2018. World Health Organization, Geneva Available at <https://apps.who.int/iris/handle/10665/274453>
  29. Sun D, Hu L, Cai Y et al (2014) A systematic review of risk factors for brain metastases and value of prophylactic cranial irradiation in non-small cell lung cancer. *Asian Pac J Cancer Prev* 15:1233–1239. <https://doi.org/10.7314/APJCP.2014.15.3.1233>
  30. Ettinger DS, Wood DE, Akerley W et al (2014) Non-small cell lung cancer, version 1.2015. *J Natl Compr Canc Netw* 12:1738–1176
  31. Lee HY, Lee KS, Kim B-T et al (2009) Diagnostic efficacy of PET/CT plus brain MR imaging for detection of extrathoracic metastases in patients with lung adenocarcinoma. *J Korean Med Sci* 24:1132. <https://doi.org/10.3346/jkms.2009.24.6.1132>
  32. Yang J, Zhang Y, Sun X, et al (2018) The prognostic value of multiorgan metastases in patients with non-small cell lung cancer and its variants: a SEER-based study. *J Cancer Res Clin Oncol* 0:0. <https://doi.org/10.1007/s00432-018-2702-9>
  33. Takenaka D, Ohno Y, Matsumoto K et al (2009) 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 FDGPET/CT, and bone scintigraphy. *J Magn Reson Imaging* 30:298–308. <https://doi.org/10.1002/jmri.21858>
  34. Cronin P, Kelly AM, Altaee D, Foerster B, Petrou M, Dwamena BA (2018) How to perform a systematic review and meta-analysis of diagnostic imaging studies. *Acad Radiol* 25:573–593. <https://doi.org/10.1016/j.acra.2017.12.007>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.