The diagnostic performance of whole-body MRI in the staging of lymphomas in adult patients compared to PET/CT and enhanced reference standard—systematic review and meta-analysis

Lukas Lambert[1]^, Andrea Burgetova[1]^, Marek Trneny[2], Bianka Bircakova[1], Jan Molinsky[2], Katerina Benesova[2], David Zogala[3], Pavel Michalek[4]

[1]Department of Radiology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic;
[2]Department of Hematology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic;
[3]Institute of Nuclear Medicine, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic;
[4]Department of Anaesthesiology, Resuscitation and Intensive Medicine, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic

Contributions: (I) Conception and design: L Lambert, M Trneny, P Michalek, A Burgetova; (II) Administrative support: P Michalek; (III) Provision of study materials or patients: L Lambert, A Burgetova; (IV) Collection and assembly of data: L Lambert, A Burgetova; (V) Data analysis and interpretation: L Lambert, A Burgetova; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Andrea Burgetova, Assoc. Prof., MD, MBA, PhD. Department of Radiology, General University Hospital in Prague, U Nemocnice 2, 128 08 Prague 2, Czech Republic. Email: andrea.burgetova@vfn.cz.

Background: Morphology highlighted by diffusion weighted imaging (DWI) is the basis of whole-body MRI (wbMRI). The aim of this study was to analyze current knowledge on the diagnostic performance of wbMRI in the pretreatment staging of patients with lymphoma.

Methods: A search for original articles reporting the diagnostic performance (sensitivity, specificity) of pretreatment (first staging or staging in relapsed patients after complete remission) wbMRI in nodal and extranodal involvement by extracranial lymphoma and the agreement of stage by the Cotswolds-modified Ann Arbor classification in adult patients compared to the reference standard (PET/CT or enhanced reference standard) was conducted in PubMed, EMBASE, Cochrane Library, ClinicalTrials.gov.

Results: Altogether 15 studies with 519 patients were included in the meta-analysis. The pooled sensitivity and specificity for nodal involvement were 0.93 (95% CI: 0.90 to 0.96) and 0.99 (95% CI: 0.98 to 1.00). For nodal staging, most studies used the size criterion of 10 mm in the short diameter (n=10) and the absence of prominent fatty hilum (n=4). Restricted diffusion on diffusion-weighted imaging as a sign of nodal involvement was either not used (n=5), used for detection (n=4), semi-quantitatively (n=4), or quantitatively (n=1). Only one study (7) relied solely on restricted diffusion as the main criterion for nodal involvement. The pooled sensitivity and specificity for extranodal involvement were 0.89 (95% CI: 0.79 to 0.98) and 0.99 (95% CI: 0.99 to 1.00). Seven studies considered diffuse splenic involvement when its long or vertical axis was greater than 13 cm regardless of the patient's physiognomy. The pooled agreement in staging (Cohen's kappa) was almost perfect (0.90, 95% CI: 0.84 to 0.95).

Discussion: The sensitivity and specificity of wbMRI in the assessment of the nodal and extranodal involvement by lymphoma is high. The agreement of wbMRI with the reference standard is almost perfect.

Keywords: Lymphoma; staging; agreement; whole-body MRI (wbMRI); PET/CT; meta-analysis

---

^ ORCID: Andrea Burgetova, 0000-0002-9975-2338; Lukas Lambert, 0000-0003-2299-4707.

Introduction

Malignant lymphoma is the sixth most common malignancy and the most common malignancy of the hematopoietic system. Lymphomas represent heterogeneous group of neoplasms with different clinical behavior from indolent to very aggressive diseases. They are traditionally divided into Hodgkin's lymphoma and non-Hodgkin lymphomas. Overall, lymphomas have 5-year survival rates between 63 and 91 percent (1,2).

Treatment of lymphomas is based on their histological (or genetic) type, stage, prognostic indices, and overall status of the patient (1). The staging considers the location and extent of the disease and the presence of clinical symptoms. The original Ann Arbor staging system was later updated by Cotswolds modification and in 2014, it was modified to reflect the widespread use of PET/CT as the Lugano classification (3). The principal groups (I to IV), remained unchanged between the original and Cotswolds-modified Ann Arbor system. In avid lymphomas, PET/CT is the workhorse for imaging of involved nodal areas and extranodal organs (2,4). In lymphomas known to have variable FDG uptake, purely morphological assessment mostly by CT or MRI is used instead (2,5).

Cross-sectional imaging has become the standard for staging, for the assessment of interim response, and after completion of chemotherapy (6). Later, patients require follow-up examinations, and some remain in watchful waiting, undergoing repeated scans. Although PET/CT is the mainstay in staging and response assessment, whole-body MRI (wbMRI) has previously been tested, whether it could be an alternative without the risk of radiation and its stochastic effects—carcinogenesis in later life (7). MRI is a safe imaging method, which is widely available, and its risks are, when observing contraindications, negligible. Therefore, it is vital to summarize current knowledge on the diagnostic performance of wbMRI in patients with lymphoma as a starting point for redefining its role, recommended imaging protocols, and interpretation of the images.

WbMRI acquires structural images that are highlighted by diffusion weighted imaging (DWI). DWI is acquired at different (low and high) b values that indicate the sensitivity to diffusion: higher b values sense smaller diffusion distances in slower moving molecules (in cellular tissues) at the expense of lower signal. Apparent diffusion coefficient (ADC), which is calculated from the decrease of signal intensity with increasing b value, is a quantifiable measure of diffusion. DWI with background body signal suppression (DWIBS) using short inversion time inversion recovery echo-planar imaging shows high signal in a tumoral mass. DWIBS allows better differentiation between reactive and diseased lymph nodes (8). The quantification of DWIBS measurements may be limited by signal decay caused by motion (in the mediastinum) and low agreement of absolute cut-off values of ADC among different studies.

The aim of this systematic review and meta-analysis was to summarize current knowledge on the diagnostic performance of pretreatment (first staging or staging in relapsed patients after complete remission) wbMRI in patients with extracranial lymphoma and the agreement of staging by the Cotswolds modification of the Ann Arbor classification in adult patients compared to the reference standard (PET/CT or enhanced reference standard).

We present the following article in accordance with the PRISMA reporting checklist (available at https://dx.doi.org/10.21037/qims-21-649).

Methods

A PRISMA checklist for Diagnostic Test Accuracy reporting was used.

A search for original articles written in the English language was performed in the following databases and registers: PubMed, EMBASE, Cochrane, clinicaltrials.gov on 12th May 2021 with the following search criteria:

PubMed

("whole body" [All fields]) and ("lymphoma" [All fields]) AND (("diffusion" [all fields]) OR ("DWI" [all fields])) AND (mri [all fields]) AND (PET [all fields]) NOT ("brain" [all fields]); 67 results.

EMBASE

'whole body' AND 'lymphoma' AND ('diffusion' OR 'dwi')
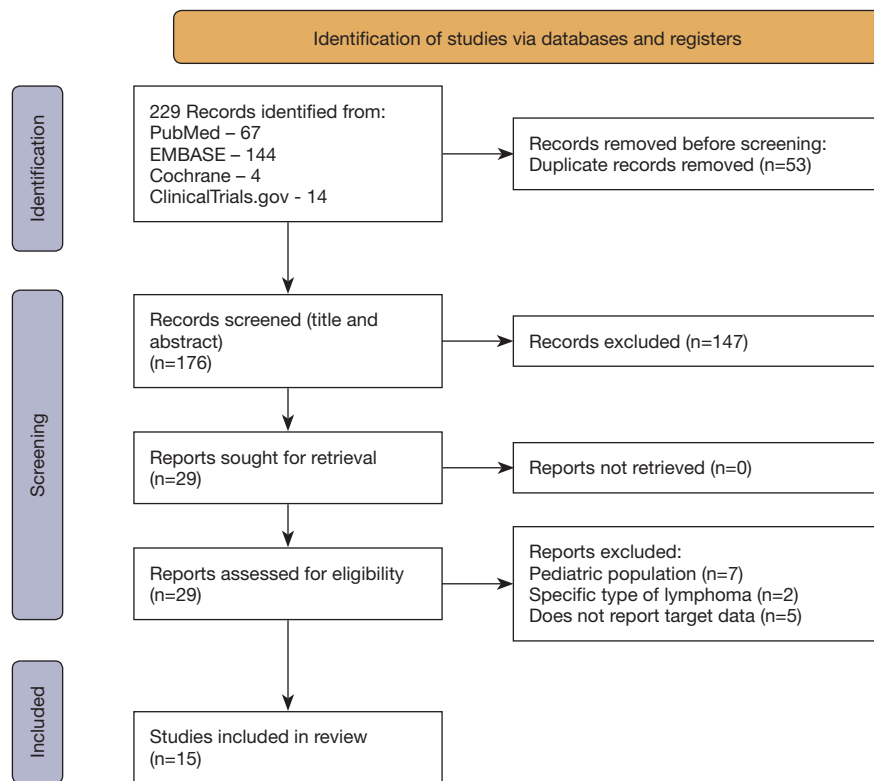
Figure 1 PRISMA 2020 flow diagram for systematic reviews.

AND 'mri' AND 'pet' NOT 'brain'; 141 results.

Cochrane Library

Whole body lymphoma MRI PET; 4 results.

ClinicalTrials.gov

Whole body lymphoma MRI PET, Adult, Older Adult; 14 studies.

Eligible studies met the following criteria: (I) adult patients; (II) pretreatment staging—newly diagnosed or relapsed after complete remission; (III) wbMRI included diffusion-weighted imaging (DWI); (IV) had a defined reference standard (PET/CT or enhanced reference standard); (V) reported sensitivity and/or specificity for nodal and/or extranodal staging and/or agreement with the reference standard; (VI) reported or allowed calculation of SE or 95% CI. We excluded studies that dealt with a primarily organ-specific type of lymphoma (primary central nervous system lymphoma, gut). The diagnosis of lymphoma and its type was based on histopathology.

The search flow is depicted in *Figure 1*. The results of the search were independently analyzed by two reviewers (LL, AB), who in case of mismatch discussed the reasons for inclusion or exclusion of an individual study. A total of 15 articles were selected for the final analysis.

The risk of bias was assessed using a QUADAS-2 tool as low, high, or unknown in consensus (*Figure 2*).

Articles included in the analysis were reviewed for inclusion criteria, patient characteristics, imaging protocols, and the reference standard used. Diagnostic performance for nodal and extranodal involvement was extracted where the data was available. Sensitivities and specificities with their 95% CIs and SEs were either retrieved or calculated from the presented data. Agreement between wbMRI and the reference standard in the stage assignment according to the Cotswolds modification of the Ann Arbor classification was calculated using Cohen's kappa statistics based on the study data or retrieved if presented with its 95% CI and SE.

Statistical analysis was performed in R (R Core Team, R foundation, Vienna, Austria). For pooling of the data, we used R package *metafor*, the inverse variance method, the DerSimonian and Laird estimator for tau^2, double

Study	Year	Risk of Bias				Applicability Concerns			
		Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard	Flow and timing
Lin	2010	+	+	+	+	+	+	+	+
Abdulqadrh	2011	+	+	+	+	+	+	+	+
Gu	2011	+	+	+	+	+	+	+	+
van Ufford	2011	+	+	+	+	+	+	+	+
stéphane	2013	+	+	+	+	+	+	+	+
Ferrari	2014	+	+	+	+	+	+	+	+
Azzedine	2015	+	+	+	?	+	+	+	?
Mayerhoefer	2015	+	+	+	+	+	+	+	+
Tsuji	2015	?	?	+	–	?	?	+	–
Albano	2016	+	+	+	?	+	+	+	?
Balbo-Musseto	2016	+	+	+	?	+	+	+	?
Gamal	2020	+	+	+	+	+	+	+	+
Kharuzhyk	2020	+	+	+	+	+	+	+	+
Latifoltojar	2020	+	+	+	+	+	+	+	+
Hong	2021	+	+	+	?	–	+	+	?

Figure 2 Risk of bias assessed using the QUADAS-2 tool. +, low; –, high; ?, unknown.

arcsine transformation of the data (Freeman-Tukey), and random effects model with Hartung-Knapp adjustment. Heterogeneity among studies was assessed using Cochran's Q and I^2 statistics. Forrest and funnel plots were generated using *forrest.meta* and *funnel* functions, respectively. Funnel plot asymmetry was calculated using Egger's test. Possible bias was also assessed by plain visual analysis of the Forrest plots and a leave-one-out analysis. A P-value below 0.05 was considered significant.

Results

All included studies originated from a single institution and apart from two (9,10), all were prospective with 15 to 140 patients (Σ519) enrolled (*Table 1*).

All wbMRI studies acquired at least two b values. The acquisition was performed in the transverse plane except for one study (22), which acquired head and neck in the

transverse plane and the rest of the body in the coronal plane. Two studies (7,11) used respiratory triggering for DWI. One study (11) involved only DWI imaging without morphological sequence.

The most commonly used morphological sequences were the STIR followed by T2 and T1 performed under free breathing, with breath-hold or respiratory-triggered. In one study (21) only, patients were given contrast material i.v. Occasionally, special acquisitions were performed for the neck and lungs (*Table 2*).

The most common risk of bias was identified in the timing of wbMRI (risk of bias in 1 study; unknown risk in 4 studies) (*Figure 2*).

For nodal staging, most studies used the size criterion of 10 mm in the short diameter (n=10) and the absence of prominent fatty hilum (n=4). Measurement of long axis was used in three studies. Restricted diffusion on diffusion-weighted imaging as a sign of nodal involvement was either

Table 1 Patient characteristics in each study included in the meta-analysis

Author	Year	Enrolment period	Study type	Age, mean ± SD [range], years	Patients	Lymphoma types		
						NHL	Aggressive NHL	HL
Lin (11)	2010	6/2008–2/2009	Prospective	48 [23–79]	15	15	15 (15 DLBCL)	0
Abdulqadhr (12)	2011	3/2008–11/2009	Prospective	47 [18–78]	31	23	18 (13 DLBCL)	8
Gu (13)	2011	11/2008–4/2010	Prospective	50±17 [20–80]	17	15	4 (3 DLBCL)	2
van Ufford (14)	2011	8/2008–10/2009	Prospective	39±8 [33–44]—HL; 61±13 [22–81]—NHL	22	20	9 (7 DLBCL)	2
Stéphane (15)	2013	6/2008–10/2009	Prospective	51 [18–84]	23	18	18 (16 DLBCL)	5
Ferrari (16)	2014	4 months	Prospective	41 [23–81]	27	14	7 (5 DLBCL)	13
Mayerhoefer (7)	2014	8/2011–1/2014	Prospective	58±16 [19–88]	140	118	34 (31 DLBCL)	22
Azzedine (17)	2015	6/2011–12/2012	Prospective	42 [22–75]	23	14	7 (7 DLBCL)	9
Tsuji (9)	2015	–	Retrospective	60±11 [36–78]	28	28	17 (17 DLBCL)	0
Albano (18)	2016	11/2013–12/2014	Prospective	42 [15–86]	68	31	16 (16 DLBCL)	37
Balbo-Mussetto (10)	2016	2/2010–5/2014	Retrospective	49 [20–76]	41	27	10 (9 DLBCL)	14
Gamal (19)	2020	5/2018–1/2020	Prospective	[16–60]	32	10	7	22
Kharuzhyk (20)	2020	2015–2018	Prospective	45±17	92	45	35 (26 DLBCL)	47
Latifoltojar (21)	2020	6/2012–11/2015	Prospective	32 [22–87]	22	8	8 (8 DLBCL)	14
Hong (22)	2021	6/2013–4/2015	Prospective	55±14 [26–82]	30	30	0	0

SD, standard deviation, NHL, non-Hodgkin lymphoma; HL, Hodgkin lymphoma; DLBCL, diffuse large B-cell lymphoma.

not used (n=5), used for detection (n=4), semi-quantitatively (n=4), or quantitatively (n=1). Only one study (7) relied solely on restricted diffusion as the main criterion for nodal involvement.

Seven studies considered diffuse splenic involvement when its long or vertical axis was greater than 13 cm and two studies (12,14) used signal intensity of the spine on DWI as a comparator for extranodal involvement (*Table 3*).

Sensitivity for nodal involvement was available in 9 studies. The pooled estimate was 0.93 (95% CI: 0.90–0.96) with low heterogeneity I^2=0.0% (0.0–64.8%); P<0.977 (*Figure 3*). Funnel plot and Egger's test did not indicate asymmetry of the distribution of the results among studies (P=0.14, Figure S1). A leave-one-out sensitivity analysis is shown on Figure S2.

Specificity for nodal involvement was available in 8 studies. The pooled estimate was 0.99 (95% CI: 0.98–1.00) with low heterogeneity I^2=0.0% (0.0–67.6%); P=1.0 (*Figure 4*). Funnel plot and Egger's test did not indicate asymmetry of the distribution of the results among studies (P=0.62, Figure S3). A leave-one-out sensitivity analysis is shown on Figure S4.

Sensitivity for extranodal involvement was available in 10 studies. The pooled estimate was 0.89 (95% CI: 0.79–0.98) with moderate heterogeneity I^2=53.0% (95% CI: 3.7–77.0%); P<0.0001 (*Figure 5*). Funnel plot and Egger's test did not indicate asymmetry of the distribution of results among studies (P=0.86, Figure S5). A leave-one-out sensitivity analysis is shown on Figure S6.

Specificity for extranodal involvement was available in 8 studies. The pooled estimate was 0.99 (95% CI: 0.99–1.00) with low heterogeneity I^2=0.0% (0.0–67.6%); P=1.00 (*Figure 6*). Funnel plot and Egger's test did not indicate asymmetry of the distribution of results among studies (P=0.67, Figure S7). A leave-one-out sensitivity analysis is shown on Figure S8.

The agreement in the stage according to the Cotswolds-modified Ann Arbor system pooled from 12 studies was 0.90 (95% CI: 0.84–0.95), with low heterogeneity of I^2=0.0% (95% CI: 0.0–58.3%), P=0.60 (*Figure 7*). Funnel plot and Egger's test did not indicate asymmetry of the distribution of results among studies (P=0.75, Figure S9). Most studies (n=11) reported agreement in the principal stages, one study (22) included E and S modifiers, where present.

Table 2 Characteristics of imaging protocols for wbMRI and PET/CT in each of the studies. Definition of the standard of reference

Author	Year	Field (T)	Manufacturer	MRI						PETCT					Reference standard
				DWI	b values (s/mm²)	STIR	T2	T1	Gd	Time (min)	MBq	CT dose	CECT	Days between PET and MRI	
Lin	2010	1.5	Siemens	TRA RT	50, 400, 800				No	30–45	5 MBq/kg	Low	No	4 (range, 1–12)	PET/CT
Abdulqadhr	2011	1.5	Philips	TRA	0, 1,000	COR RT	TRA RT (lesions)	COR BH	No	50	5 MBq/kg	Std	No	6 pts: same day; 25 pts: 7.4±6.6; 9.3 months	ERS - imaging, biopsy, clinical staging, FU (6–14 months)
Gu	2011	3	Philips	TRA	0, 1,000	TRA	TRA		No	48±4	4.8 MBq/kg	Std	No	2±1 (≤1 week)	PET/CT
van Ufford	2011	1.5	Siemens	TRA	0, 1,000	COR		COR	No	45–55	3 MBq/kg	Low	No	within 15 days	PET/CT
Stéphane	2013	1.5	Philips	TRA	0, 1,000	yes (ns)		yes (ns)	No	40	5 MBq/kg	Low	No	3.5 (range, 1–9)	PET/CT
Ferrari	2014	1.5	Philips	TRA	0, 500	COR (BH)		COR (BH)	No	20–25	370–550 MBq (3.7 MBq/kg)	Low	No	Within 10 days	PET/CT
Azzedine	2015	3	Philips	TRA	0, 1,000		COR		No	44	5 MBq/kg	Std	No	–	Nodal involvement – PET/CTStaging - ERS – imaging, physical examination, bone marrow aspiration and biopsy
		1.5		TRA	0, 1,000		COR (TrueFisp)		No	12[1]					
Mayerhoefer	2015	3	Siemens	TRA (RT)	50, 400, 1,000			TRA	No	20–30	270–340 MBq (by BW)	Std	Yes	≤7 days	ERS - imaging, biopsy, follow-up
Tsuji	2015	3	GE	TRA	0, 1,000	COR	COR		No	35	185 MBq	Low	Yes	–	PET/CT
Albano	2016	1.5	Philips	TRA	0, 800	COR (RT)		COR (BH)	No	23–29	3.7 MBq/kg	Low	No	–	PET/CT
Balbo-Mussetto	2016	1.5	Philips	TRA	0, 1,000	COR	COR	COR	No	45	BW-related dose	–	No	–	ERS - biopsy, PET/CT, clinical FU, histology, imaging FU
Gamal	2020	1.5	Philips	TRA	0, 1,000	COR		COR	No	20	5.5 MBq/kg	Std	No	Within 10 days	ERS - imaging, histology, FU 10 months
Kharuzhyk	2020	1.5	GE	TRA	0, 800	COR	TRA RT (chest)	TRA	No	~42	291±49 MBq	Std	No	7±11	ERS - imaging, biopsy, FU 6 months
Latifoltojar	2020	3	Philips	TRA	0, 100, 300, 1,000	COR	COR	COR	20 mL @3 mL/s²	~40	5.5 MBq/kg	Std	No	10 (range, 0–44)	ERS – PET/CT, FU
Hong	2021	3	Philips	TRA, COR	0, 800	TRA, COR (RT)		COR	No	26–28	–	–	–	–	ERS - imaging, clinical staging, FU imaging (116±65 days), histopathology, patients' clinical course

¹, this reported value is likely erroneous; ², Gadoterate meglumine (Dotarem®). DWI, diffusion-weighted imaging; STIR, short-tau inversion recovery; RT, respiratory trigger; BH, breath-hold; TrueFisp, steady state balanced sequence (Siemens); CECT, contrast-enhanced CT; std, standard; TRA, transverse plane; COR, coronal plane; ERS, enhanced reference standard; FU, follow-up; ~, rough estimate; –, data not available; BW, body weight; ns, not specified.

A single study (21) that reported agreement based on the Lugano criteria, was not included. A leave-one-out sensitivity analysis is shown on Figure S10.

Altogether, the sensitivity analysis did not reveal any significant bias with any of the included studies.

Discussion

In this meta-analysis, we summarized the diagnostic performance of wbMRI in the pretreatment assessment of nodal and extranodal involvement by lymphoma and its agreement with FDG PET/CT or enhanced reference standard (including FDG PET/CT) in the staging. We showed that the pooled sensitivity for nodal and extranodal involvement is 0.93 and 0.89 while the specificity is nearly 1.0. The agreement in staging (Cohen's kappa) is almost perfect.

Diffusion-weighted imaging with background suppression (DWIBS) is the mainstay of whole-body MR imaging not only in lymphoma but also in other cancers (23,24). WbMRI can be reliably performed both at 1.5T and 3T scanners that provide sufficient field homogeneity. The suppression of background signal is performed by inversion pulse to null the signal of adipose tissue. DWI measures the restriction of random Brownian motion that occurs in tissues with a high cellular component where diffusion is hindered by cellular membranes (25). In the analyzed studies, wbMRI was performed during free breathing in most of them. Respiratory triggering used in two studies (7,11) results in a better depiction of mediastinal and hilar lymph nodes. However, the scan time is usually prolonged by a factor of 2.5. FDG-PET provides different metrics which is a glycolytic activity (25).

For nodal staging, most studies used the size criterion of 10 mm in the short diameter and the absence of prominent fatty hilum. Smaller lymph nodes may be considered if they exhibit other malignancy features including the absence of fatty hilum, local grouping, atypical localization, topographical relationship with involved organs, or show highly restricted diffusion compared to other structures at the same level—usually, the spinal cord was chosen as the comparator (20). Inguinal, iliac, and axillary lymph nodes frequently present with restricted diffusion, so in these regions, attention must be paid to exclude fatty hilum in the measurement of their short axis. It is important to notice that from the studies reporting the sensitivity for nodal involvement <0.9, there were two relying both on the morphological criteria and semiquantitative assessment of

Table 3 Comparison of criteria for nodal and extranodal involvement among the included studies

Author	Year	Nodal involvement		Extranodal involvement			Splenic involvement	
		Morphology	DWI	Morphology	DWI		Morphology	DWI
Lin	2010	SA >10 mm	ADC SI < SI muscle	SI abnormalities or mass lesions (excluding benign eg. cysts)	Abnormal SI on ADC		Diameter >13 cm, focal abnormality	
Abdulqadhr	2011	SA >10 mm	×	T1 and STIR: area of abnormal SI	DWI SI > spinal cord; focal increase in SI[1]		T1 and STIR or DWI: focal abnormal signal intensity	
Gu	2011	LA >10 mm, SI > surroundings on STIR or SI < surroundings on T2; clustered small lymph nodes; unusual location; central necrosis	DWI SI > SI spinal cord	T2 and STIR: lesions were detected by identification of a mass lesion or signal abnormalities	×		T2 and STIR: lesions were detected by identification of a mass lesion or signal abnormalities	
van Ufford	2011	SA >10 mm	Detection	STIR and T1: abnormal SI, mass	DWI SI > SI spinal cord; focal increase in SI[1]		Area of abnormal SI on T1 and STIR, focal restricted diffusion	

Table 3 (*continued*)

Table 3 (*continued*)

Author	Year	Nodal involvement		Extranodal involvement		Splenic involvement
		Morphology	DWI	Morphology	DWI	
Stéphane	2013	SA >10 mm SA >15 mm for axillary and femoral except LNN with fatty hilum and thin cortex	Visual analysis (except axillary and femoral LNN), ADC<0.75×10^{-3} mm^2/s	x	Diffusion signal abnormality	Longest diameter >13 cm (without liver cirrhosis)
Ferrari	2014	SA >10 mm	Detection	T1 or STIR: altered signal	DWI SI > surrounding tissues	Altered signal in T1 or STIR, DWI SI > surrounding tissues
Azzedine	2015	LA >10 mm	DWI SI > SI of the muscle on the same image slice	x	Areas of restricted diffusion	Diameter >13 cm; areas of abnormal SI on DWI
Mayerhoefer	2015	x	restricted diffusion	x	Restricted diffusion	SI inhomogeneity or well-circumscribed lesions with restricted diffusion
Tsuji	2015	–	–	–	–	–
Albano	2016	LA >15 mm (excl. fatty hilum and thin cortex)	ADC <0.8×10^{-3} mm^2/s	Focal lesions >1cm, signal abnormalities	Areas of restricted diffusion	Focal lesion or max. diameter >13 (excl. in cirrhosis)
Balbo-Mussetto	2016	SA >10 mm: (neck and mediastinum SA >15: abdomen unusual location, necrotic	DWI SI > SI spinal cord	x	Abnormal SI relative to surrounding tissue; focally increased SI[1]	Focally increased SI
Gamal	2020	SA >10 mm	x	T1 or STIR: altered signal	DWIBS SI > SI surrounding tissues	Altered signal in T1 or STIR, DWI SI > surrounding tissues (kidney), max. diameter > 13 cm
Kharuzhyk	2020	SA >10 mm (excl. fatty hilum, thin parenchyma); SA ≤10 mm: round shape, the absence of hilum, with local grouping, and localization in atypical regions	x	Areas of pathological SI	x	Vertical size >13 cm
Latifoltojar	2020	SA >10 mm	x	Specific for each organ: abnormal moderate-high SI infiltration, abnormal focal SI	Specific for each organ: abnormal moderate-high SI infiltration, abnormal focal SI	Moderate-high SI infiltration in continuum with an adjacent mass; low SI (relative to spleen) discrete foci within spleen
Hong	2021	SA >10 mm, loss of fatty hilum, heterogeneity	x	STIR: focal or diffuse abnormal SI; asymmetrical enlargement of paired organs	Focal increase in DWI SI[1]	Focal or diffuse abnormal SI on STIR, max. diameter >13 cm

[1], in tissues with normally impeded diffusion. SI, signal intensity; DWI, diffusion-weighted imaging; ADC, apparent diffusion coefficient; x, not used; −, not reported; SA, short axis; LA, long axis; T2, T2 weighted sequence; STIR, short tau inversion recovery; T1, T1 weighted sequence.
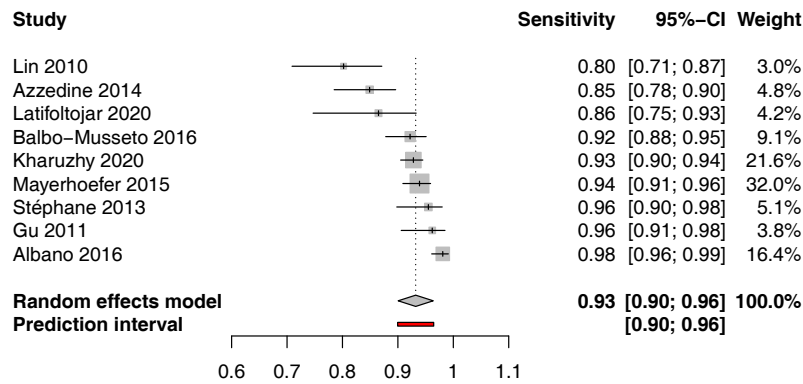
Figure 3 Forrest plot of the sensitivity of wbMRI in the detection of nodal involvement by lymphoma. 95% CI, 95% confidence interval.
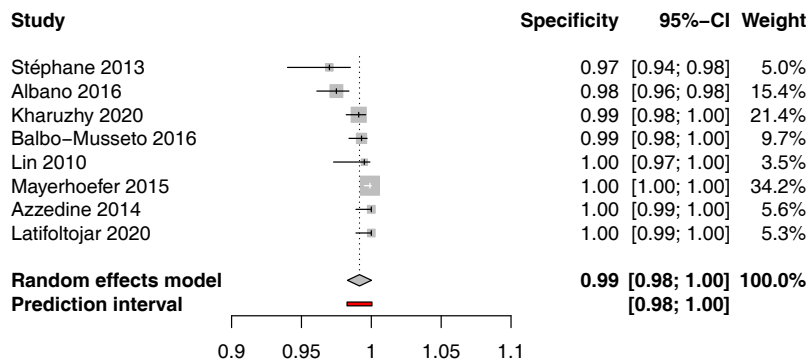

Figure 4 Forrest plot of the specificity of wbMRI in the detection of nodal involvement by lymphoma. 95% CI, 95% confidence interval.
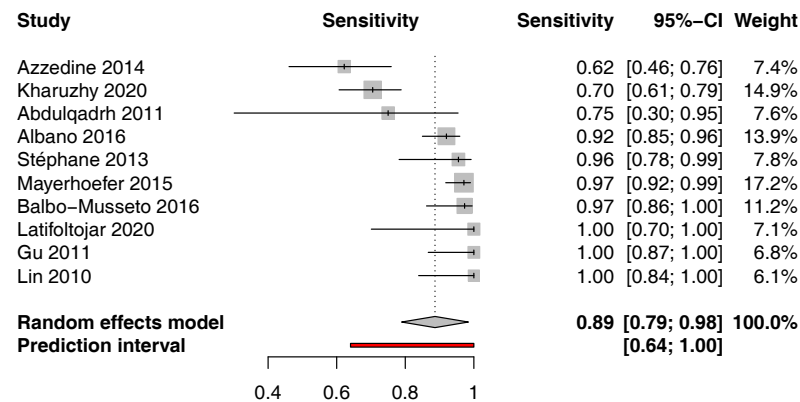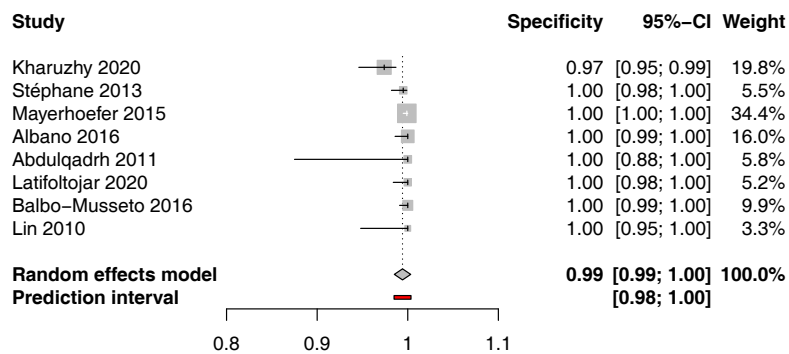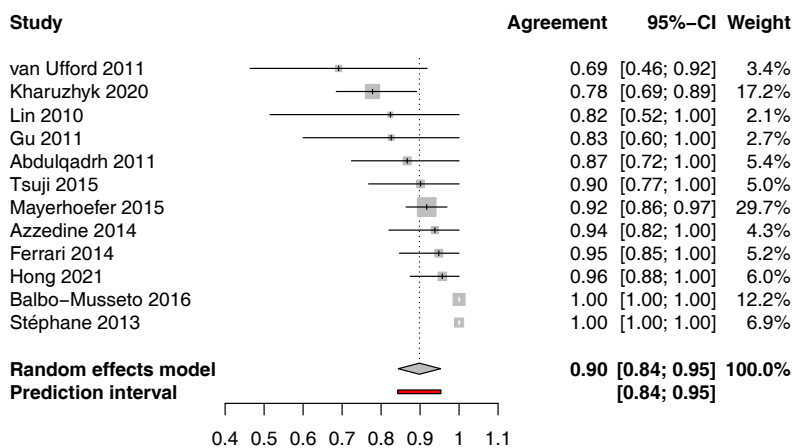

Figure 5 Forrest plot of the sensitivity of wbMRI in the detection of extranodal involvement by lymphoma. 95% CI, 95% confidence interval.

Figure 6 Forrest plot of the specificity of wbMRI in the detection of extranodal involvement by lymphoma. 95% CI, 95% confidence interval.

Figure 7 Forrest plot of the agreement of wbMRI and the reference standard in the stage. 95% CI, 95% confidence interval.

DWI or ADC (11,17).

Extranodal involvement was based on focal changes on T1, T2, or STIR of non-benign nature. Focal DWI increase compared to the surrounding tissue is required in organs that normally exhibit restricted diffusion. To assess diffuse infiltration, signal intensity on DWI can also be visually compared to other organs at the same level such the spinal cord or other surrounding tissues. The signal intensity on DWIBS of a single organ is different between adjacent stations because of corrections of field homogeneity and shimming. The diagnostic performance of wbMRI is limited in the detection of small pulmonary infiltrates and diffuse splenic involvement. Seven studies considered diffuse splenic involvement when its long or vertical axis was greater than 13 cm regardless of the patient's physiognomy. The normal size of the spleen is significantly influenced by body weight and sex and its upper 95% confidence limits

exceed 13 cm in men taller than 170 cm (26). Twelve studies included SI abnormities on morphological and/or DWI sequences.

In all but one study, DWI with background suppression was performed without respiratory compensation which prolongs the examination. Because of respiratory and cardiac motion, DWI is less efficient in the evaluation of hilar, mediastinal, or pulmonary involvement. Most studies performed DWI imaging in the axial plane, only one in the coronal plane. Imaging in the transverse plane is obviously more suitable for the construction of fusion images with transverse anatomical acquisitions such as STIR, T2, or T1 that resemble that of PET/CT. According to our experience, fast STIR and T2 sequences can be performed even without respiratory compensation while T1 requires breath-hold at the thoracic and abdominal stations.

So why has not wbMRI become the preferred method

for pretherapeutic lymphoma staging? Kharuzhyk *et al.* have shown that wbMRI and PET/CT in the nodal and extranodal involvement have comparable diagnostic accuracy (20). We believe that the main considerations are the higher complexity and time demand for wbMRI. The mean time needed to perform wbMRI was 35±15 min. This requires that one hour is reserved for an examination which is from the reimbursement point of view equal to an examination that requires half an hour. Another advantage of PET/CT is the benefit of functional information in the standardized assessment of the treatment response, which has been firmly and clearly established in the form of the Deauville criteria (27). We believe that there is still unexplored potential of wbMRI in the quantification of the treatment response using T1 and T2 mapping, susceptibility weighted imaging, and even fat fraction (13).

Study limitations. A major limitation of this meta-analysis is the disproportion in lymphoma types considered (*Figure 2*). Secondly, there was heterogeneity in MRI imaging protocols and MRI scanners (field strength, manufacturer). Thirdly, stage agreement was based on the Cotswolds-modified Ann Arbor classification because only one study reported using the more recent Lugano criteria. Lastly, some studies used only PET/CT as the reference standard while the majority used enhanced reference standard, which included histology, and follow-up examinations.

In conclusion, the sensitivity of wbMRI in the assessment of the nodal and extranodal involvement by lymphoma is above 0.93 and 0.89 while the specificity is nearly 1.0. Agreement in the staging is almost perfect with Cohen's kappa of 0.9. Further studies are needed to develop more accurate criteria for the involvement of lymph nodes especially in difficult areas such as in the mediastinum and pulmonary hili. Also, the criteria for diffuse splenic involvement require the attention of further studies.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the PRISMA reporting checklist. Available at https://dx.doi.org/10.21037/qims-21-649

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/qims-21-649). LL serves as an unpaid editorial board member of *Quantitative Imaging in Medicine and Surgery*. LL and AB report that this study received an institutional support from the Ministry of Health of the Czech Republic (MH CZ-DRO, General University Hospital in Prague - VFN, 00064165). The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Lewis WD, Lilly S, Jones KL. Lymphoma: Diagnosis and Treatment. Am Fam Physician 2020;101:34-41.
2. McCarten KM, Nadel HR, Shulkin BL, Cho SY. Imaging for diagnosis, staging and response assessment of Hodgkin lymphoma and non-Hodgkin lymphoma. Pediatr Radiol 2019;49:1545-64.
3. Cheson BD, Fisher RI, Barrington SF, Cavalli F, Schwartz LH, Zucca E, et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J Clin Oncol 2014;32:3059-68.
4. Meignan M, Hutchings M, Schwartz LH. Imaging in Lymphoma: The Key Role of Fluorodeoxyglucose-Positron Emission Tomography. Oncologist 2015;20:890-5.
5. Alessandrino F, DiPiro PJ, Jagannathan JP, Babina G, Krajewski KM, Ramaiya NH, Giardino AA. Multimodality imaging of indolent B cell lymphoma from diagnosis to transformation: what every radiologist should know. Insights Imaging 2019;10:25.
6. Johnson SA, Kumar A, Matasar MJ, Schöder H,

Rademaker J. Imaging for Staging and Response Assessment in Lymphoma. Radiology 2015;276:323-38.

7. Mayerhoefer ME, Karanikas G, Kletter K, Prosch H, Kiesewetter B, Skrabs C, Porpaczy E, Weber M, Pinker-Domenig K, Berzaczy D, Hoffmann M, Sillaber C, Jaeger U, Müllauer L, Simonitsch-Klupp I, Dolak W, Gaiger A, Ubl P, Lukas J, Raderer M. Evaluation of diffusion-weighted MRI for pretherapeutic assessment and staging of lymphoma: results of a prospective study in 140 patients. Clin Cancer Res 2014;20:2984-93.

8. Donners R, Yiin RSZ, Koh DM, De Paepe K, Chau I, Chua S, Blackledge MD. Whole-body diffusion-weighted MRI in lymphoma-comparison of global apparent diffusion coefficient histogram parameters for differentiation of diseased nodes of lymphoma patients from normal lymph nodes of healthy individuals. Quant Imaging Med Surg 2021;11:3549-61.

9. Tsuji K, Kishi S, Tsuchida T, Yamauchi T, Ikegaya S, Urasaki Y, Fujiwara Y, Ueda T, Okazawa H, Kimura H. Evaluation of staging and early response to chemotherapy with whole-body diffusion-weighted MRI in malignant lymphoma patients: A comparison with FDG-PET/CT. J Magn Reson Imaging 2015;41:1601-7.

10. Balbo-Mussetto A, Cirillo S, Bruna R, Gueli A, Saviolo C, Petracchini M, Fornari A, Lario CV, Gottardi D, De Crescenzo A, Tarella C. Whole-body MRI with diffusion-weighted imaging: a valuable alternative to contrast-enhanced CT for initial staging of aggressive lymphoma. Clin Radiol 2016;71:271-9.

11. Lin C, Luciani A, Itti E, El-Gnaoui T, Vignaud A, Beaussart P, Lin SJ, Belhadj K, Brugières P, Evangelista E, Haioun C, Meignan M, Rahmouni A. Whole-body diffusion-weighted magnetic resonance imaging with apparent diffusion coefficient mapping for staging patients with diffuse large B-cell lymphoma. Eur Radiol 2010;20:2027-38.

12. Abdulqadhr G, Molin D, Aström G, Suurküla M, Johansson L, Hagberg H, Ahlström H. Whole-body diffusion-weighted imaging compared with FDG-PET/CT in staging of lymphoma patients. Acta Radiol 2011;52:173-80.

13. Gu J, Chan T, Zhang J, Leung AY, Kwong YL, Khong PL. Whole-body diffusion-weighted imaging: the added value to whole-body MRI at initial diagnosis of lymphoma. AJR Am J Roentgenol 2011;197:W384-91.

14. van Ufford HM, Kwee TC, Beek FJ, van Leeuwen MS, Takahara T, Fijnheer R, Nievelstein RA, de Klerk JM. 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;196:662-9.

15. Stéphane V, Samuel B, Vincent D, Joelle G, Remy P, Francois GG, Jean-Pierre T. Comparison of PET-CT and magnetic resonance diffusion weighted imaging with body suppression (DWIBS) for initial staging of malignant lymphomas. Eur J Radiol 2013;82:2011-7.

16. Ferrari C, Minoia C, Asabella AN, Nicoletti A, Altini C, Antonica F, Ficco M, Guarini A, Maggialetti N, Rubini G. Whole body magnetic resonance with diffusion weighted sequence with body signal suppression compared to (18) F-FDG PET/CT in newly diagnosed lymphoma. Hell J Nucl Med 2014;17 Suppl 1:40-9.

17. Azzedine B, Kahina MB, Dimitri P, Christophe P, Alain D, Claude M. Whole-body diffusion-weighted MRI for staging lymphoma at 3.0T: comparative study with MR imaging at 1.5T. Clin Imaging 2015;39:104-9.

18. Albano D, Patti C, La Grutta L, Agnello F, Grassedonio E, Mulè A, Cannizzaro G, Ficola U, Lagalla R, Midiri M, Galia M. Comparison between whole-body MRI with diffusion-weighted imaging and PET/CT in staging newly diagnosed FDG-avid lymphomas. Eur J Radiol 2016;85:313-8.

19. Gamal GH. Whole-body magnetic resonance/diffusion-weighted sequence with background signal suppression (WB-MR/DWIBS) vs. 18F-FDG PET/CT in diagnosis of lymphoma. Egyptian Journal of Radiology and Nuclear Medicine 2020;51:210.

20. Kharuzhyk S, Zhavrid E, Dziuban A, Sukolinskaja E, Kalenik O. Comparison of whole-body MRI with diffusion-weighted imaging and PET/CT in lymphoma staging. Eur Radiol 2020;30:3915-23.

21. Latifoltojar A, Duncan MKJ, Klusmann M, Sidhu H, Bainbridge A, Neriman D, Fraioli F, Lambert J, Ardeshna KM, Punwani S. Whole Body 3.0 T Magnetic Resonance Imaging in Lymphomas: Comparison of Different Sequence Combinations for Staging Hodgkin's and Diffuse Large B Cell Lymphomas. J Pers Med 2020;10:284.

22. Hong GS, Chae EJ, Ryu JS, Chae SY, Lee HS, Yoon DH, Suh C. Assessment of naive indolent lymphoma using whole-body diffusion-weighted imaging and T2-weighted MRI: results of a prospective study in 30 patients. Cancer Imaging 2021;21:5.

23. Michielsen K, Dresen R, Vanslembrouck R, De Keyzer F, Amant F, Mussen E, Leunen K, Berteloot P, Moerman P, Vergote I, Vandecaveye V. Diagnostic value of whole body diffusion-weighted MRI compared to computed

tomography for pre-operative assessment of patients suspected for ovarian cancer. Eur J Cancer 2017;83:88-98.

24. Wang D, Huo Y, Chen S, Wang H, Ding Y, Zhu X, Ma C. Whole-body MRI versus 18F-FDG PET/CT for pretherapeutic assessment and staging of lymphoma: a meta-analysis. Onco Targets Ther 2018;11:3597-608.

25. Stecco A, Buemi F, Iannessi A, Carriero A, Gallamini A. Current concepts in tumor imaging with whole-body MRI with diffusion imaging (WB-MRI-DWI) in multiple myeloma and lymphoma. Leuk Lymphoma 2018;59:2546-56.

26. Chow KU, Luxembourg B, Seifried E, Bonig H. Spleen Size Is Significantly Influenced by Body Height and Sex: Establishment of Normal Values for Spleen Size at US with a Cohort of 1200 Healthy Individuals. Radiology 2016;279:306-13.

27. Voltin CA, Mettler J, Grosse J, Dietlein M, Baues C, Schmitz C, Borchmann P, Kobe C, Hellwig D. FDG-PET Imaging for Hodgkin and Diffuse Large B-Cell Lymphoma-An Updated Overview. Cancers (Basel) 2020;12:601.
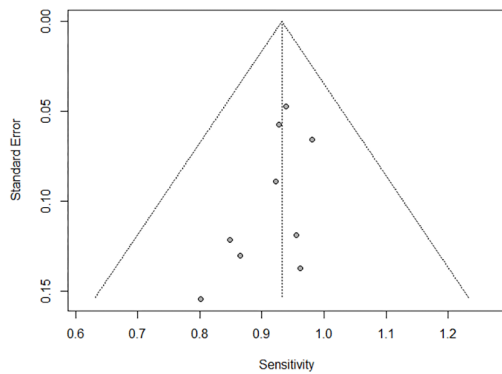
Figure S1 Funnel plot of the distribution of the sensitivity of wbMRI in the detection of nodal involvement by lymphoma.
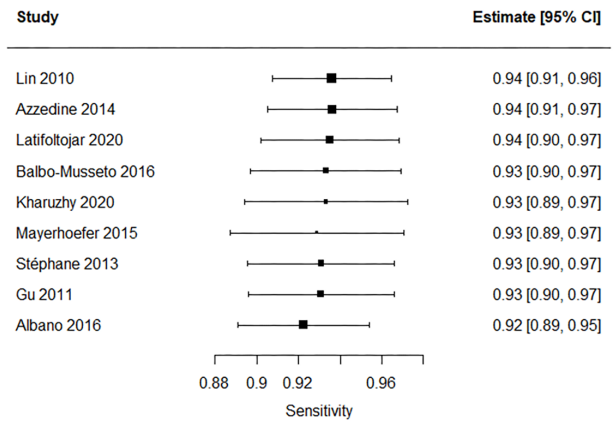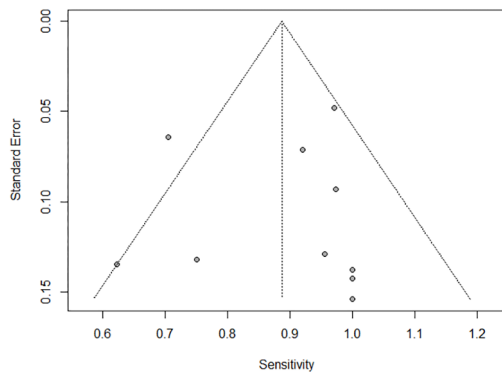

Figure S2 Sensitivity of wbMRI in the detection of nodal involvement by lymphoma shown on a leave-one-out sensitivity analysis.
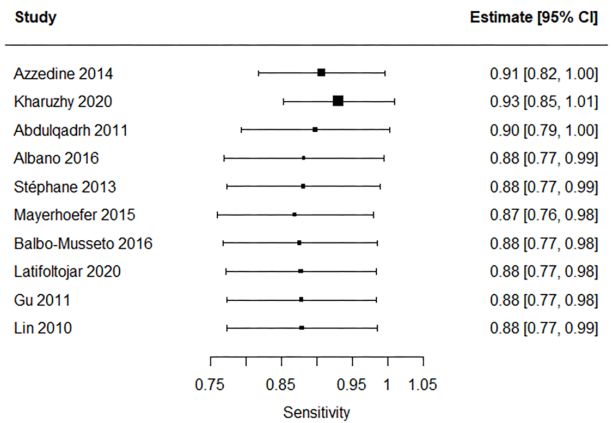

Figure S3 Funnel plot of the distribution of the specificity of wbMRI in the detection of nodal involvement by lymphoma.
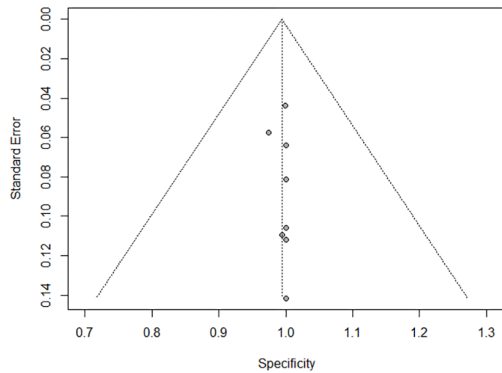

Figure S4 Specificity of wbMRI in the detection of nodal involvement by lymphoma shown on a leave-one-out sensitivity analysis.
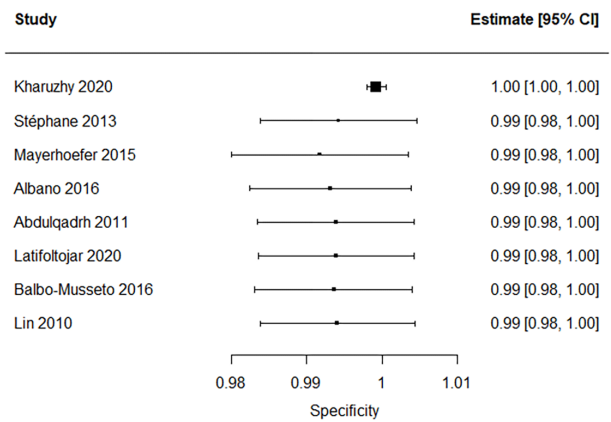
Figure S5 Funnel plot of the distribution of the sensitivity of wbMRI in the detection of extranodal involvement by lymphoma.


Figure S6 Sensitivity of wbMRI in the detection of extranodal involvement by lymphoma shown on a leave-one-out sensitivity analysis.


Figure S7 Funnel plot of the distribution of the specificity of wbMRI in the detection of extranodal involvement by lymphoma.

Figure S8 Specificity of wbMRI in the detection of extranodal involvement by lymphoma shown on a leave-one-out sensitivity analysis.
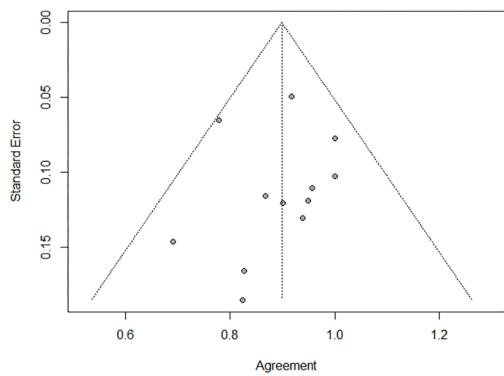
Figure S9 Funnel plot of the distribution of the agreement of wbMRI and the reference standard in the stage.
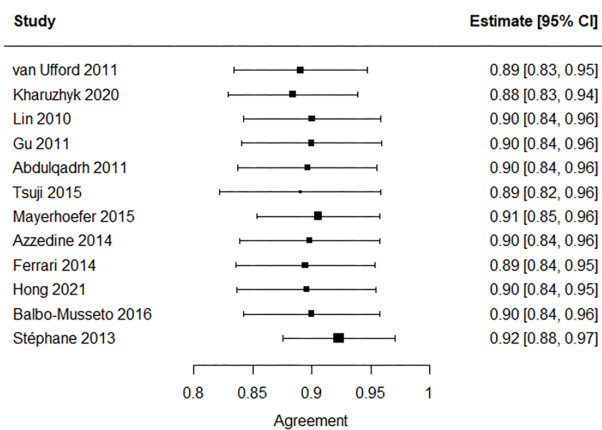
Figure S10 Agreement of wbMRI and the reference standard in the stage shown on a leave-one-out sensitivity analysis.