EJC Paediatric Oncology 1 (2023) 100016

Contents lists available at ScienceDirect

EJC Paediatric Oncology

journal homepage: www.journals.elsevier.com/ejc-paediatric-oncology

A meta-analysis on the diagnostic performance of whole-body MRI for the initial staging of Hodgkin lymphoma in children and adults using FDG-PET/CT as a reference standard

Suzanne Spijkers (M.D., Ph.D.)[a,*], Annemieke S. Littooij (M.D., Ph.D.)[a,b], Auke Beishuizen (M.D., Ph.D.)[b], Marnix G.E.H. Lam (M.D., Ph.D.)[a,b], Rutger A.J. Nievelstein (M.D., Ph.D.)[a,b]

[a] Department of Radiology and Nuclear Medicine, University Medical Centre Utrecht/Wilhelmina Children's Hospital, Utrecht University, Utrecht, the Netherlands
[b] Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Hodgkin disease
Magnetic resonance imaging
Imaging, whole body
Staging, neoplasm
Meta-analysis

ABSTRACT

Background: Staging of Hodgkin lymphoma is important for determining prognosis and treatment planning. The current gold standard is FDG-PET/CT, but WB-MRI could be a radiation free alternative.
Objective: A meta-analysis of all published data on the diagnostic performance of WB-MRI for the initial staging of Hodgkin lymphoma using FDG-PET/CT as a reference standard.
Evidence Acquisition: Both the PubMed/MEDLINE and EMBASE databases were systematically searched (updated until March 14, 2023) for studies that compared WB-MRI with FDG-PET/CT for staging Hodgkin lymphoma. The "quality assessment of diagnostic accuracy studies" tool (QUADAS-2) was used to assess methodological quality. Pooled staging accuracy, sensitivity and specificity of WB-MRI compared to FDG-PET/CT was calculated for determining stage and for both nodal and extra-nodal staging. A sensitivity analysis for children and adults was performed.
Evidence Synthesis: A total of nine studies with a combined total of 297 Hodgkin lymphoma patients were included. Pooled sensitivity and specificity for nodal staging were 94% (95%CI 0.92–0.96) and 99% (95%CI 0.98–1.00) respectively. For extra-nodal staging sensitivity and specificity were 90% (95%CI 0.74–0.96) and 100% (95%CI 0.99–1.00). For disease stage, the pooled accuracy was 92% for pediatric studies (95%CI 0.86–0.96), 94% for adult studies (95%CI 0.87–0.97) and 92% (95%CI 0.87–0.96) for all studies combined.
Conclusion: When using FDG-PET/CT as a reference standard, WB-MRI shows high sensitivity and specificity for both nodal and extra-nodal staging and for determining disease stage both in children and adults.
Clinical Impact: WB-MRI could be used as a good radiation-free alternative for FDG-PET/CT in Hodgkin lymphoma staging.

1. Introduction

The incidence of Hodgkin lymphoma varies with age, gender and country and accounts for 0.4% of new malignancies worldwide [1]. In adolescents, Hodgkin lymphoma is the most common malignancy and accounts for 13% of malignancies [2]. Accurate staging at diagnosis is of great importance for determining prognosis and treatment planning. Currently, a wide range of imaging modalities - including magnetic resonance imaging (MRI), contrast-enhanced computed tomography (CE-CT), ultrasound and conventional radiographs - are used worldwide for staging. The current gold standard is ^{18}F-fluorodeoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) [3–5].

Due to the combination of both functional and anatomical information and the high FDG avidity of Hodgkin lymphomas, FDG-PET/CT has high sensitivity and specificity for staging [6–8]. Thus, FDG-PET/CT is considered the current best available imaging modality.

Abbreviations: CE-CT, Contrast-enhanced computed tomography; CI, Confidence interval; DWI, Diffusion weighted imaging; FDG-PET/CT, ^{18}F-fluorodeoxyglucose positron emission tomography /computed tomography; WB-MRI, Whole body magnetic resonance imaging

* Correspondence to: University Medical Centre Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands.
E-mail address: suzannespijkers@outlook.com (S. Spijkers).

https://doi.org/10.1016/j.ejcped.2023.100016
Received 22 March 2023; Received in revised form 9 May 2023; Accepted 31 May 2023
Available online 5 June 2023

However, during staging, treatment evaluation and follow-up patients receive multiple imaging examinations resulting in repeated exposure to ionizing radiation. The amount of administered radiation varies and depends on local protocols, the type of scanner, age and the use of either a high-dose or a low-dose CT. However, although FDG-PET/CT radiation doses have declined with advancing technology, large differences in administered dosages are still present between hospitals [9]. Especially in children and adolescents, this repeated exposure is not desirable because of their increased vulnerability to radiation and the years they have ahead in which secondary radiation induced malignancies can develop [10–16]. Secondary tumor risk due to follow-up CT scans in patients with lymphoma has been established as well as the remarkable increase of cancer risk associated with the use of PET/CT [17–19]. According to Brenner and Hall, up to 2% of malignancies are related to CT scans [20].

In recent literature, whole-body MRI (WB-MRI) has been investigated as a radiation-reduced alternative to FDG-PET/CT for the initial staging of Hodgkin lymphoma [21–27]. An important limitation of WB-MRI is that the detection of involved nodal sites still relies on size criteria. However, a major step forward was the possibility of acquiring functional information as well using diffusion weighted imaging (DWI), which was added to the conventional WB-MRI protocols [28].

Malignant lymphomas form a diverse group of tumors with a variety of treatments, different FDG-avidities and patient populations and although there are few review studies investigating WB-MRI for staging, all of them included multiple types of lymphoma, which leads to heterogeneous studies [29,30]. Furthermore, no separate analyses for children have been performed, although children and adolescents are one of the main age groups in which Hodgkin lymphoma is prevalent. Therefore, the aim of this study was to systematically compare studies, with a special focus on studies that included children, on the diagnostic accuracy of WB-MRI for the initial staging of Hodgkin lymphoma using FDG-PET/CT as a reference standard.

2. Materials and methods

2.1. Literature search

The PubMed/MEDLINE and EMBASE databases were comprehensively searched for studies in English comparing WB-MRI with an FDG-PET/CT-based reference standard for the initial staging of Hodgkin lymphoma. No date limits were used. The search was updated until March 14, 2023. A combination of the following search terms and their synonyms was used: Hodgkin lymphoma, WB-MRI, diffusion weighted imaging, FDG-PET/CT and staging. The full search strategy is shown in supplementary table 1.

2.2. Study selection and data extraction

All diagnostic test accuracy studies written in English in which separate Hodgkin lymphoma analyses were performed for comparing WB-MRI and FDG-PET/CT for initial staging were included. No language, age or sample size limits were used. Studies were excluded if: there were no separate Hodgkin lymphoma data available (e.g., for studies in which multiple lymphoma types were assessed), there were overlapping study populations, and/or the study did not assess full staging (e.g. studies only mentioning bone marrow involvement).

After performing the search (supplementary table 1), all titles and abstracts were screened for eligibility. For potentially eligible studies the full text articles were obtained. Based on the in- and exclusion criteria a final selection was made. As a final step in the search strategy, bibliographies of eligible studies were screened for potential additional references. Authors from lymphoma studies in which no separate Hodgkin lymphoma data was provided were contacted for the raw study data regarding the included Hodgkin lymphoma patients to possibly be able to evaluate the separate Hodgkin lymphoma data.

The following data were extracted from all eligible studies: the first author, year of publication, publication type/study design, number of participants and their demographic data (gender, mean age and range). Furthermore, data on the exact reference standard, WB-MRI protocols and image interpretation methods were extracted. For the study outcome, all data concerning Hodgkin lymphoma staging (both nodal, extra-nodal and full stage) were extracted. Data extraction was performed by one reviewer (S.S.) and confirmed by an independent second reviewer (A.S.L.).

2.3. Quality assessment

The quality (risk of bias and applicability) of the included studies was assessed by two independent reviewers (S.S. and A.S.L.) using the Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2) [31]. The risk of bias was assessed in each of the following domains: patient selection, index test, reference standard and flow and timing. The first three domains were also assessed in terms of concerns regarding applicability. Risks were described as ‘low’, ‘high’ or ‘unclear’. Differences between the reviewers were solved in consensus.

2.4. Statistical analysis

All analyses were performed using R statistical software (R version 3.4.1., R Foundation for Statistical Computing, Vienna, Austria). Forest plots for sensitivity and specificity including 95% confidence intervals (CI) were made for the extracted nodal and extra-nodal staging data (lesion-based analysis). A forest plot for the accuracy of WB-MRI to determine full stage was constructed as well (patient-based analysis). Staging accuracy was defined as the proportion of patients who were staged correctly by WB-MRI as compared to the reference standard. Since limited disease (stages I and II) needs less treatment than advanced disease (stages III and IV) an additional analysis concerning the percentages of upstaging and downstaging by WB-MRI was performed with the dichotomized variables (limited disease and advanced disease). Wilson score CI’s were calculated before pooling the estimates [32].

Statistical heterogeneity was assessed by assessing the forest plots visually and by using the Higgins I^2 statistics (measuring inconsistency across studies) [33]. I^2 values above 50% were interpreted as substantial heterogeneity.

To estimate pooled sensitivity, specificity and accuracy a random-effects generalized mixed model was used using the meta package in R. A sensitivity analyses was performed for the separate adult and pediatric studies for staging accuracy.

3. Results

3.1. Study selection

The search in both PubMed/MEDLINE and EMBASE resulted in 463 articles after the removal of duplicates. All titles and abstracts were screened and 27 articles that met the inclusion criteria based on title and abstract were selected for full-text screening. Six articles directly met the inclusion criteria and 15 additional articles were found eligible if the authors could provide us with separate data for their included Hodgkin lymphoma patients. After contacting all corresponding authors of these latter 15 articles, we were able to include three more articles, thus a total of nine articles could be included for meta-analysis. Fig. 1 shows the article inclusion flow chart.

In Table 1 the main characteristics of the nine included studies are presented. These studies comprised a total of 297 Hodgkin lymphoma patients [21–27,34,35]. All studies used WB-MRI as index test. For the reference standard, most studies used an FDG-PET/CT-based reference standard that included other clinical and imaging findings [21–24,27,34,35]. Two studies used the FDG-PET/CT findings only as reference standard [25,26]. Detailed imaging comparisons are provided

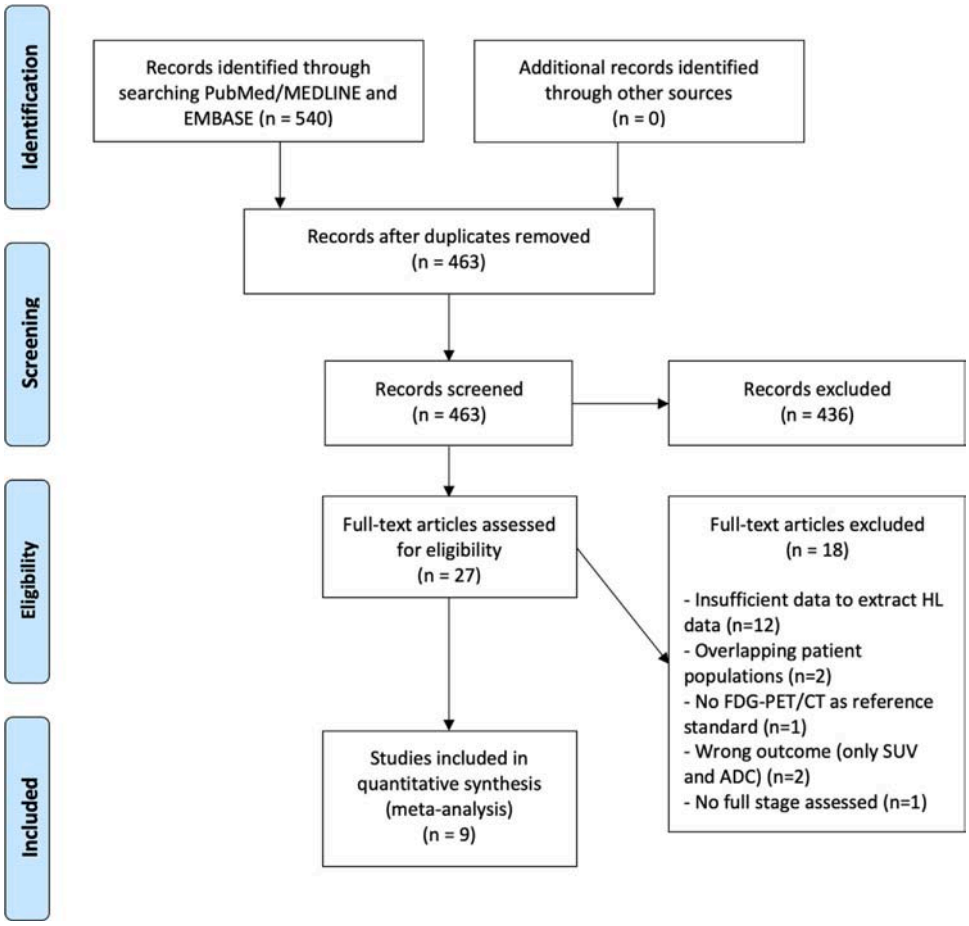


Fig. 1. Flowchart for selection and exclusion of studies.

in Table 2. All studies reported Ann-Arbor staging accuracy of WB-MRI and for six out of nine studies sensitivity and specificity for nodal and extra-nodal staging could be calculated from the provided 2 × 2 tables. One study did report sensitivity and specificity for nodal and extra-nodal staging, but authors did not report the amount of true and false positives and negatives [26]. Table 3 shows the definitions used by the included studies for positive findings for both nodal and extra nodal disease and describes the use of DWI (if applicable).

3.2. Quality assessment

Fig. 2 summarizes the risk of bias concerns and applicability judgments for all included studies. The graph shows percentages of the included studies for which either low, high or unclear was scored on each item.

For the risk of bias concerns, the index test was scored low in all studies. The reference standard was scored high in one study, since the reference standard was formed by one radiologist, after FDG-PET/CT reading by the nuclear medicine physicians [35]. Flow and timing was scored unclear in two studies, since the time interval between the index test (WB-MRI) and the reference standard (FDG-PET/CT) was not mentioned in the article [23,25]. Patient selection was scored unclear in two studies. In one study it was unclear why three patients were excluded [22] and in the other study it was not clear whether or not patients were prospectively or retrospectively included, neither was it mentioned if the patients were consecutively selected [24].

In the assessment of the concerns regarding applicability both patient selection and the reference standard were scored low in all studies.

Table 1
Main characteristics of the included studies.

Author	Year	Study type	Hodgkin lymphoma patients (n)	Mean age and range (years)	Sex (Male/Female)	Adult/pediatric
Albano et al.	2016	Prospective	37	31 (15–57)	21/16	Mixed
Balbo-Mussetto et al.	2016	Unclear	14	40 (20–65)	7/7	Adult
Ferrari et al.	2014	Prospective	13	41 (23–81)	5/8	Adult
Kharuzhyk et al.	2020	Prospective	47	35 (19–59)	20/27	Adult
Latifoltojar et al.	2019	Prospective	50	16* (6–19)	32/18	Pediatric
Littooij et al.	2014	Prospective	23	15 (6–21)	10/13	Pediatric
Mayerhoefer et al.	2014	Prospective	22	NA	NA	Adult
Punwani et al.	2010	Prospective	21	NA	NA	Pediatric
Spijkers et al.	2020	Prospective	68	14 (7–17)	33/35	Pediatric

* median. NA = not applicable, data not provided.

Table 2
Detailed comparison imaging techniques between included studies.

Author, year	Index test (WB-MRI)				Reference standard				
	Field strength	Sequence	Extent	Average total scan time	Type	Extent	CT	Dose^{18}FDG	Average total acquisition time
Albano et al. [23]	1.5 T	T1, T2 STIR, DWI	Craniocaudal coverage of 185.5 cm, (= full body for most patients)	30 min	FDG-PET/CT	Skull base to pelvis	NA	3.7 MBq/kg	25–30 min
Balbo-Mussetto et al. [24]	1.5 T	T1, T2 TSE, T2 STIR, DWI	Head/neck, chest, abdomen, pelvis, and femurs	45 min	FDG-PET/CT combined with bone marrow biopsy, clinical stage and follow-up data	Skull base to pelvis	CE-CT	NA	NA
Ferrari et al., [25]	1.5 T	T1, T2 STIR, DWI	Head/neck, thorax, abdomen, and pelvis	20–25 min	FDG-PET/CT	Head to pelvis	Low-dose	3.7MBq/kg	20 min
Kharuzhyk et al. [35]	1.5 T	T1, T2 STIR, DWI	Upper orbital edge to middle third femur	40 min	FDG-PET/CT combined with biopsy data, imaging results and follow-up data	NA	NA	NA	NA
Latifoltojar et al. [22]	1.5 T	T1, T2, DWI	Skull to mid-thigh	60 min	Enhanced reference standard, FDG-PET/CT, expert panel, clinical outcomes, imaging results and follow-up data	NA	Non-contrast CT	14–370 MBq	NA
Littooij et al. [21]	1.5 T	T1, T2 STIR, DWI	Head/neck, chest, abdomen, and pelvis	50–60 min	FDG-PET/CT-based, with expert panel, clinical outcomes and imaging results	Skull base to midfemur	Low-dose or CE-CT	5.18–7.4 MBq/kg	30 min
Mayerhoefer et al. [26]	3 T	T1, DWI	Vertex to upper thigh	NA	FDG-PET/CT and histology results	NA	CE-CT	300 MBq	NA
Punwani et al. [34]		T2	Neck, chest, abdomen, and pelvis	25–30 min	FDG-PET/CT-based, with expert panel, clinical outcomes and imaging results	Skull base to midthigh	CE-CT	370 MBq	NA
Spijkers et al. [27]	1.5 T	T1, T2 STIR, DWI	Head/neck, chest, abdomen, and pelvis	50–60 min	FDG-PET/CT-based, with expert panel, clinical outcomes and imaging results	Skull base to midthigh	Low-dose	2 MBq/kg	30 min

CE-CT = contrast-enhanced computed tomography, DWIBS = diffusion weighted imaging, FDG-PET/CT = 18 F-fluorodeoxyglucose positron emission tomography /computed tomography, MBq = mega Becquerel, STIR = short tau inversion recovery, T = Tesla, TSE = turbo spin echo, WB-MRI = whole-body magnetic resonance imaging, NA = not applicable, not mentioned in the article.

Table 3
Definitions for nodal and extra-nodal involvement at WB-MRI and the use of DWI.

Author, year	Nodal disease	Extra nodal disease*	DWI
Albano et al. [23]	Long axis diameter (> 1,5 cm) combined with ADC value cutoff (< 0.8 × _10 −3mm2/s)	Spleen: focal lesions or splenomegaly (longest diameter > 13 cm) without cirrhosis Extra-nodal lesions: were detected identifying focal lesions (> 1 cm), signal abnormalities and areas of restricted diffusion	A combination of ADC cutoff value and visual assessment (*b*0 and *b*800)
Balbo-Mussetto et al. [24]	Short axis diameter (> 1 cm in neck and mediastinum and 1,5 cm in the abdomen)	Every area of abnormal signal intensity relative to the surrounding tissue was considered pathological. For tissues with normally impeded diffusion (including spleen) any focally increased signal intensity was considered positive for tumor involvement	Visual assessment of *b*0 and *b*1000
Ferrari et al. [25]	Short axis diameter (> 1 cm)	For extra nodal regions or spleen any areas with altered signal in T1w or STIR, showing signal intensity in DWIBS higher than surrounding tissues, have been considered positive for lymphoma localization	Visual assessment of *b*0–500 and *b*1000
Kharuzhyk et al. [35]	Short axis diameter (> 1 cm)	Spleen: vertical size > 13 cm was considered diffuse involvement Extra nodal lesions: foci or areas of pathological signal intensity	Visual assessment of *b*0 and *b*800
Latifoltojar et al. [22]	Long axis diameter > 2 cm or lymph nodes 1–2 cm with ADC ≤ 1.2 or lymph nodes < 1 cm with ADC ≤ 0.8	Extra-nodal lesions: were detected identifying focal lesions, signal abnormalities and areas of restricted diffusion	A combination of visual assessment and ADC cutoff values. *b*0, *b*100, *b*300, *b*500.
Littooij et al. [21]	Short axis diameter (> 1 cm)	Spleen: discrete nodules or enlargement (splenic index > 725) Extra-nodal lesions: were detected identifying focal lesions (> 1 cm), signal abnormalities and areas of restricted diffusion	Visual assessment of *b*0 and *b*1000
Mayerhoefer et al. [26]	Lymph nodes were rated as positive if they had a long-axis diameter > 1.5 cm, or a long axis and short-axis diameter of each > 1 cm.	Spleen: signal inhomogeneity or well-circumscribed lesions with restricted diffusion. Extra nodal lesions: positive for lymphoma if restricted diffusion was seen on DWI	Visual assessment of *b*50, *b1000* and ADC map
Punwani et al. [34]	Short axis diameter (> 1 cm)	Spleen: low signal discrete foci within spleen discrete from any adjacent lymphatic mass Extra-nodal lesions: signal abnormalities, discrete foci	NA, no DWI was used
Spijkers et al. [27]	Short axis diameter (> 1 cm)	Spleen: discrete nodules or enlargement Extra-nodal lesions: were detected identifying focal lesions (> 1 cm), signal abnormalities and areas of restricted diffusion	Visual assessment of *b*0, *b*100, and *b*800

NA = not applicable, data not provided. DWI = Diffusion weighted imaging. ADC= apparent diffusion coefficient.
* Often described in more detail in the original articles.

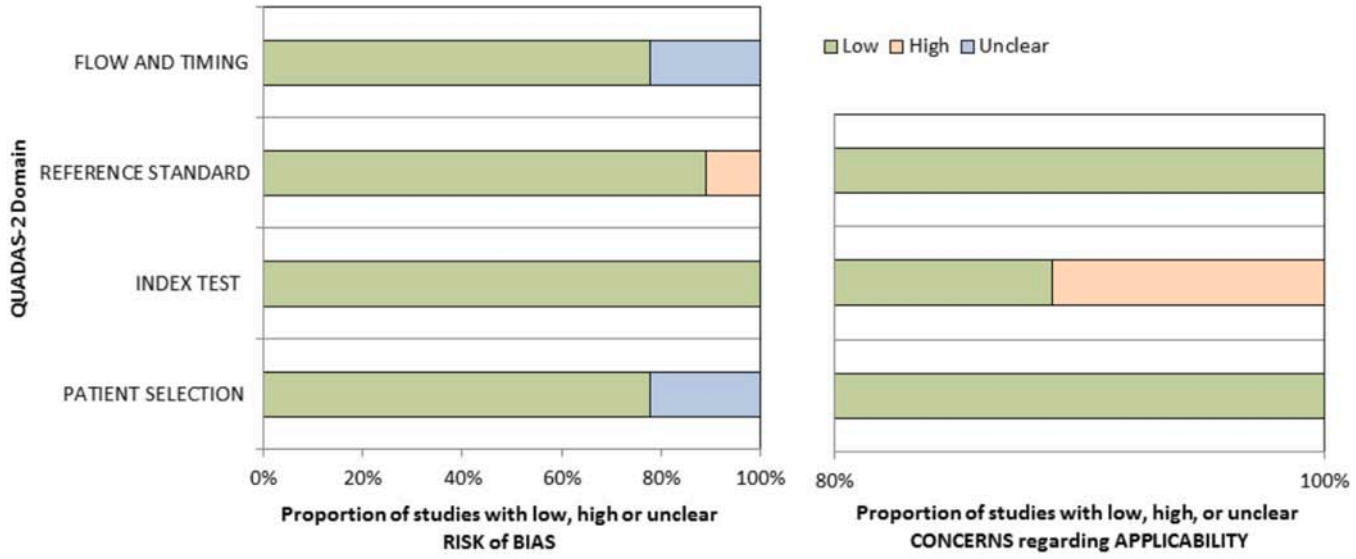


Fig. 2. Risk of bias and applicability concerns. Quality assessment of diagnostic accuracy studies (QUADAS-2) scores for each category are shown. The left part of the image shows the reviewers' concerns (low, high or unclear) regarding risk of bias across four domains: flow and timing, the reference standard, the index test and patient selection. The right part of the image shows the proportion of studies with concerns (low, high or unclear) about the applicability of the studies for the research question of this meta-analysis. The applicability concerns are scored for three domains: the reference standard, the index test and patient selection.

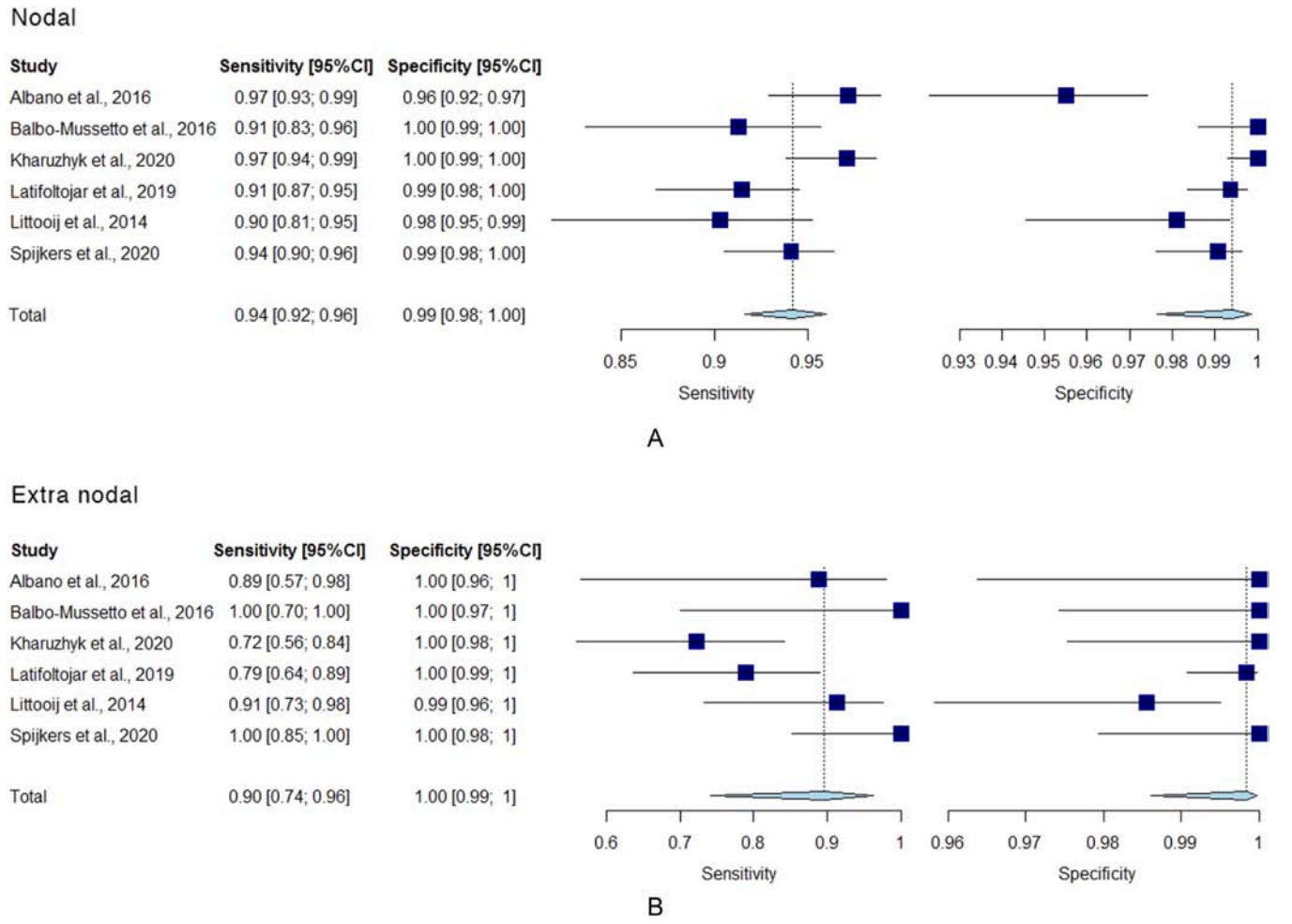


Fig. 3. Forest plot. Nodal (a) and extra-nodal (b) staging sensitivity and specificity of WB-MRI calculated against an FDG-PET/CT-based reference standard.

For the index test, one study was scored high due to the absence of DWI in the WB-MRI protocol [34].

3.3. Nodal and extra-nodal staging

Sensitivity and specificity of WB-MRI for detecting both nodal and extra-nodal disease were calculated against the FDG-PET/CT-based reference standards for all six articles that reported the number of true and false positive and negative sites [21–24,27,35]. Forest plots are presented in Fig. 3. Together, the studies comprised a combined total of 3250 nodal and 1539 extra-nodal sites.

Heterogeneity between studies was the highest for nodal specificity (I^2 = 86%), and heterogeneity was also substantial for both nodal and extra-nodal sensitivity and extra-nodal specificity (I^2 = 49%, 63% and 51% respectively).

For nodal disease, sensitivity ranged from 0.90 to 0.97 whereas specificity ranged from 0.96 to 1.00. Pooled summary estimates for sensitivity and specificity were 0.94 (95%CI 0.92–0.96) and 0.99 (95%CI 0.98–1.00) respectively for nodal disease detection.

For the detection of extra-nodal disease by WB-MRI, sensitivity and specificity in the included studies ranged from 0.72 to 1.00 and 0.99–1.00 respectively. Pooled sensitivity of WB-MRI for detecting extra-nodal disease was 0.90 (95%CI 0.74–0.96). For specificity the pooled estimate was 1.00 (0.99–1.00).

Three studies had to be excluded for nodal and extra-nodal sensitivity and specificity analysis since authors did not report the necessary data [25,26]. Mayerhoefer et al. did provide sensitivity and specificity for nodal (88% and 100% respectively) and extra-nodal (both 100%) staging, but no raw data was provided [26]. In Table 4, all sensitivities and specificities are shown.

3.4. Staging accuracy WB-MRI

In Fig. 4 **(a)**, the forest plot for staging accuracy of WB-MRI versus the reference standard is shown. Heterogeneity between studies was low with I^2 = 33%. Balbo-Mussetto et al. reported the highest staging accuracy (1.00, 95%CI 0.78–1.00) whereas staging accuracy was the lowest in the study of Latifoltojar et al. (0.82, 95%CI 0.69–0.90) and Ferrari et al. (0.85, 95%CI 0.58–0.96) [22,24,25]. The pooled staging accuracy was 0.92 (95%CI 0.87–0.96).

A separate analysis for adult and pediatric studies was performed as well (Fig. 4 **(b)**). Heterogeneity (I^2) between studies was 0% for adult studies [24–26,35] and 48% for pediatric studies [21,22,27]. Pooled staging accuracies were comparable for both adult and pediatric studies: 0.94 (95%CI 0.87–0.97) for the adult studies and 0.92 (95%CI 0.86–0.96) for pediatric studies.

In Fig. 5 two forest plots are shown to summarize the percentages of clinically relevant (defined as having implications for treatment) upstaging and downstaging by WB-MRI as compared to the reference standards for each study. Heterogeneity (I^2) between studies was 0% for upstaging and 25% for downstaging. Pooled results showed a 3% (95%CI 1–6%) upstaging rate across studies and a 2% (95%CI 1–7%) downstaging rate when comparing WB-MRI with the reference standard.

Table 4
Nodal and extra-nodal staging and staging accuracy of WB-MRI against the reference standard.

Author, year	Nodal staging sensitivity	Nodal staging specificity	Extra-nodal staging sensitivity	Extra-nodal staging specificity	Ann Arbor staging accuracy
Albano et al. [23]	97%	96%	89%	100%	95%
Balbo-Mussetto et al. [24]	91%	100%	100%	100%	100%
Ferrari et al. [25]	NA	NA	NA	NA	85%
Kharuzhyk et al. [35]	97%	100%	72%	100%	96%
Latifoltojar et al. [22]	91%	99%	79%	100%	82%
Littooij et al. [21]	90%	98%	91%	99%	87%
Mayerhoefer et al. [26]	88%	100%	100%	100%	91%
Punwani et al. [34]	NA	NA	NA	NA	91%
Spijkers et al. [27]	94%	99%	100%	100%	97%

Values are either calculated from data provided by authors or in the article or presented in the article. NA = not applicable, data not provided.

4. Discussion

This meta-analysis included nine studies comparing WB-MRI to the current reference standard FDG-PET/CT for staging Hodgkin lymphoma. In comparison with FDG-PET/CT, WB-MRI shows high sensitivity and specificity for both nodal and extra-nodal staging and for determining disease stage. To the best of our knowledge, this meta-analysis is the first to systematically assess WB-MRI for staging Hodgkin lymphoma in both adults and children.

Although our main aim was to focus on the results in children, we chose to include all studies that compared WB-MRI and FDG-PET/CT in Hodgkin lymphoma staging. Given the age distribution of Hodgkin lymphoma showing two peaks – late childhood to early adulthood and late adulthood; the best approach would probably be to include children and adolescents in studies. Since most studies focus on either children or adults (or a combination) we chose to include all of them and perform separate analyses.

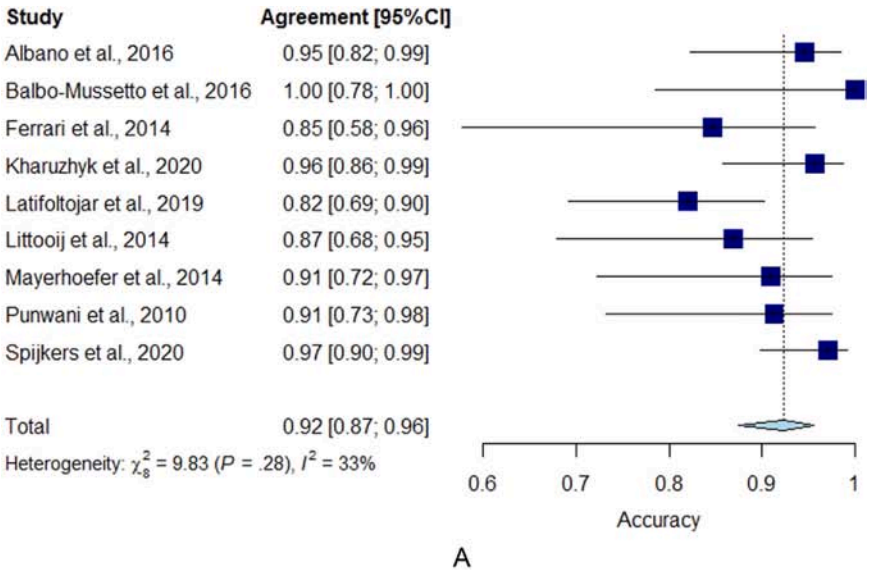
Both the imaging techniques that are used as index test and reference standard for this meta-analysis – WB-MRI and FDG-PET/CT respectively – have favorable and less favorable characteristics. FDG-PET/CT is the established reference standard, is widely used and useful for follow-up imaging as well. However, disadvantages include the use of ionizing radiation and, in younger children, the need for sedation. WB-MRI on the other hand is a radiation free alternative that provides excellent anatomical detail. Drawbacks are less availability, the need for more experienced readers and the need for sedation in children under the age of six. Furthermore, nodal disease in normal sized lymph nodes can be difficult to detect at WB-MRI due to the fact that normal lymph nodes show restricted diffusion as well.

All included studies reported high sensitivity and specificity for initial staging of Hodgkin lymphoma. This was the case for both lesion-based analyses and the patient-based staging analyses. Sensitivity was found to be somewhat lower than specificity for all analyses, but was still 90% or higher for both nodal and extra-nodal staging. For extra-nodal staging, the sensitivity of WB-MRI was the lowest (90%), detecting extra-nodal disease is thus more difficult with WB-MRI compared to the reference standard. Specificity approached 100% for all studies for both nodal and extra-nodal staging. For the patient-based analysis, the staging accuracy, all studies reported accuracies between 82% and 100%. Pooled staging accuracy was 92% for all studies combined, 94% for adult studies and 92% for pediatric studies. WB-MRI is thus a highly sensitive, specific and accurate staging method for Hodgkin lymphoma. In terms of clinical relevance, upstaging and downstaging by WB-MRI compared to the reference standard occurred in only 2–3% (pooled estimates). These are the percentages of patients in which staging by WB-MRI would have led to different treatment planning.

Although these results are promising, for the wider interpretation the question rises whether WB-MRI can keep up with FDG-PET/CT for response assessment and follow-up. Current literature shows that WB-MRI (including visual assessment of DWI) is not yet accurate enough for response assessment due to the difficulty of distinguishing rest lesions from active lesions [22,36–42]. Further research could focus on the question whether or not a baseline FDG-PET/CT is needed for accurate response assessment by FDG-PET/CT, or that it can be replaced by WB-MRI only at baseline. [22,36–42].

The included studies had a few limitations. In general, the included studies were homogeneous enough for meta-analysis and Higgins I^2 remained well below 50%. However, some heterogeneity was present between studies, especially for the lesion-based analyses. This was mainly caused by the different imaging protocols of both the index test and reference standard (Table 2) and the differences in assessing disease presence across studies (Table 3). These differences in both imaging protocols and interpretation of the images need to be harmonized across institutions in the future in order to provide higher levels of evidence, as was also stated by the Oncology Task Force of the European Society of Radiology [43]. Furthermore, a limitation of the included studies is the lack of a true gold standard as a reference standard. Although a histopathological diagnosis for all disease sites would be the most reliable evidence of Hodgkin lymphoma involvement, invasive exploration of all potential disease sites is of course ethically undesirable. Seven of the nine studies used a form of FDG-PET/CT-based enhanced reference standard including clinical outcomes, histopathological outcomes and other imaging modalities as well [21,22,24,26,27,34]. The remaining two studies used only FDG-PET/CT as a reference standard [23,25].

There are a few limitations of this meta-analysis that need to be addressed. First, a total of 12 studies in which Hodgkin lymphoma patients were analyzed together with other lymphoma types needed to be excluded because no additional information was provided by the corresponding author. This may have caused selection bias to some extent. Second, the different WB-MRI imaging protocols and FDG-PET/CT protocols used in the included studies may have caused heterogeneity (Table 2), However, the statistical heterogeneity was still less than 50% for almost all analysis. Third, this meta-analysis did not include studies in which only specific disease locations were assessed [44–48]. Although those publications contribute to location-based knowledge for staging Hodgkin lymphoma, no conclusions based on full disease stage were made. To provide a homogeneous result and study outcome, for this meta-analysis only studies assessing full disease stage



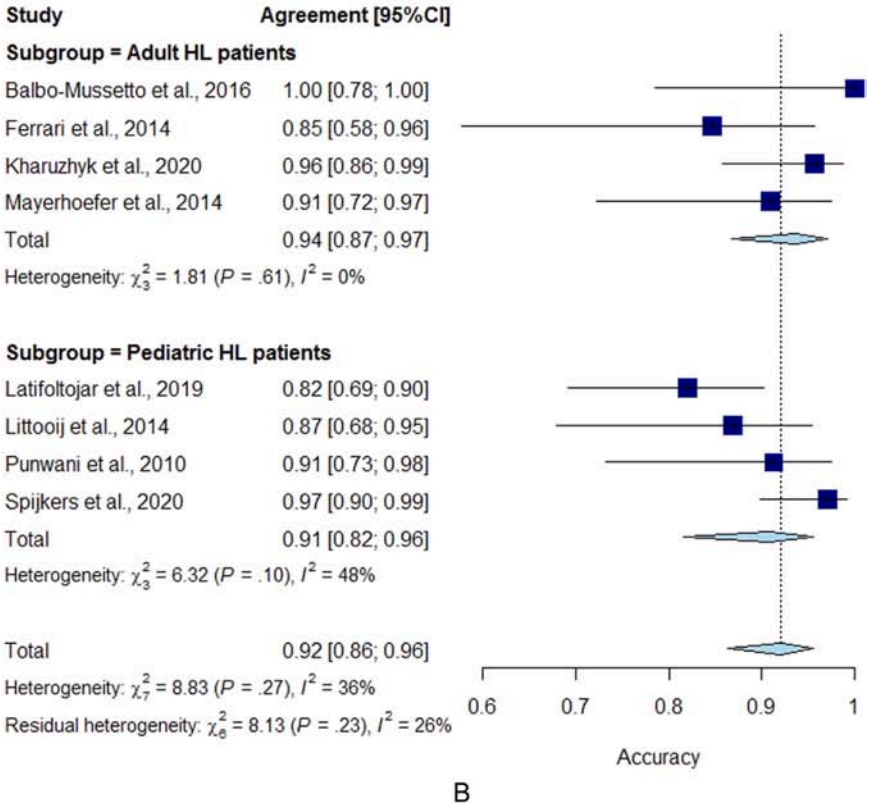


Fig. 4. Forest plots. Staging accuracy of WB-MRI. (a) all studies combined. (b) sensitivity analyses, adult studies and pediatric studies.

were included. And fourth, since mainly MRI studies with positive results were found, publication bias may have been present. The young adults (up to 35 years-old) are now grouped in the adult group instead of in the 'young adult group' that would follow the natural distribution of Hodgkin lymphoma. And finally, due to the relatively low incidence of Hodgkin lymphoma, only nine studies could be included and the included studies all comprised limited numbers of patients.

To conclude, WB-MRI has shown to be a highly sensitive and specific imaging method for initial staging of Hodgkin lymphoma. Excellent agreement with the reference standard was seen for both nodal and extra-nodal staging and for determining disease stage in children and adults. Therefore, WB-MRI may be a viable radiation free alternative for FDG-PET/CT in staging Hodgkin lymphoma for both children and adults.

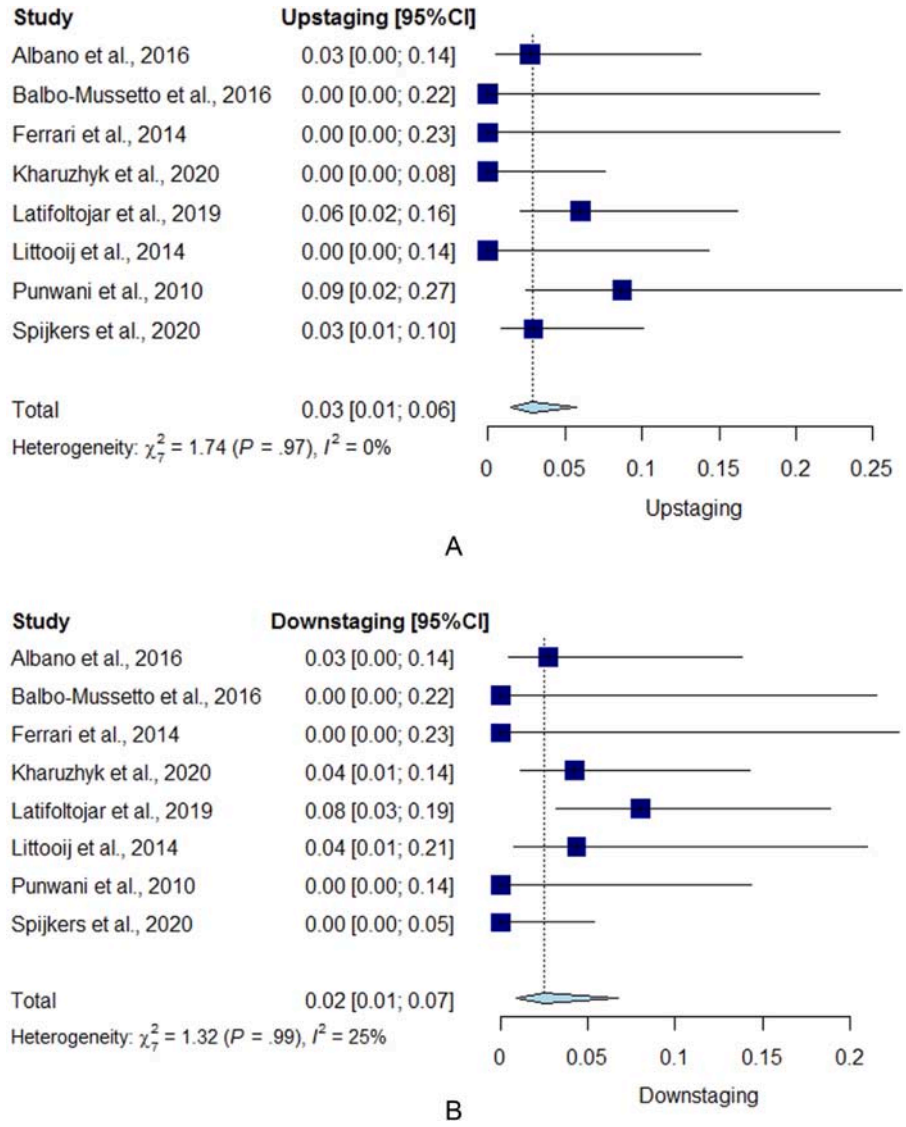


Fig. 5. Forest plots. Percentages of clinically relevant (i.e. implications for treatment planning) upstaging and downstaging by WB-MRI compared to the reference standard. (a) Upstaging. (b) Downstaging.

Funding

This project was financially supported by the Children Cancer Free Foundation (KiKa, project number 87). The data collection, analysis and interpretation of data, the writing of the paper and the decision to submit were not influenced by the Children Cancer Free Foundation.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: R.A.J. Nievelstein reports financial support was provided by the Children Cancer Free Foundation.

Acknowledgements

We would like to thank dr. Domenico Albano, dr. Annalisa Balbo-Mussetto and dr. Siarhei Kharuzhyk for providing additional study data.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejcped.2023.100016.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (2018) 394–424, https://doi.org/10.3322/caac.21492

[2] American Cancer Society, Cancer Facts & Figures 2020, American Cancer Society, Atlanta, 2020.

[3] S.F. Barrington, N.G. Mikhaeel, L. Kostakoglu, M. Meignan, M. Hutchings, S.P. Müeller, L.H. Schwartz, E. Zucca, R.I. Fisher, J. Trotman, O.S. Hoekstra, R.J. Hicks, M.J. O'Doherty, R. Hustinx, A. Biggi, B.D. Cheson, Role of imaging in the staging and response assessment of lymphoma: Consensus of the international conference on malignant lymphomas imaging working group, J. Clin. Oncol. 32 (2014) 3048–3058, https://doi.org/10.1200/JCO.2013.53.5229

[4] B.D. Cheson, R.I. Fisher, S.F. Barrington, F. Cavalli, L.H. Schwartz, E. Zucca, T.A. Lister, Recommendations for initial evaluation, staging, and response assessment of hodgkin and non-hodgkin lymphoma: the lugano classification, J. Clin. Oncol. 32 (2014) 3059–3067, https://doi.org/10.1200/JCO.2013.54.8800

[5] R. Kluge, L. Kurch, T. Georgi, M. Metzger, Current role of FDG-PET in pediatric Hodgkin's lymphoma, Semin. Nucl. Med. 47 (2017) 242–257, https://doi.org/10.1053/j.semnuclmed.2017.01.001

[6] C. La Fougère, W. Hundt, N. Bröckel, T. Pfluger, A. Haug, B. Scher, M. Hacker, K. Hahn, M. Reiser, R. Tiling, Value of PET/CT versus PET and CT performed as separate investigations in patients with Hodgkin's disease and non-Hodgkin's lymphoma, Eur. J. Nucl. Med. Mol. Imaging 33 (2006) 1417–1425, https://doi.org/10.1007/s00259-006-0171-x

[7] T.C. Kwee, R.M. Kwee, R.A.J. Nievelstein, Imaging in staging of malignant lymphoma: a systematic review, Blood 111 (2008) 504–516, https://doi.org/10.1182/blood-2007-07-101899

[8] L. Uslu, J. Donig, M. Link, J. Rosenberg, A. Quon, H.E. Daldrup-Link, Value of18F-FDG PET and PET/CT for evaluation of pediatric malignancies, J. Nucl. Med 56 (2015) 274–286, https://doi.org/10.2967/jnumed.114.146290

[9] J.F. Schäfer, S. Gatidis, H. Schmidt, B. Gückel, I. Bezrukov, C.A. Pfannenberg, M. Reimold, M. Ebinger, J. Fuchs, C.D. Claussen, N.F. Schwenzer, Simultaneous Whole-Body PET/MR Imaging in Comparison to PET/CT in Pediatric Oncology: Initial Results, Radiology 273 (2014) 220–231, https://doi.org/10.1148/radiol.14131732

[10] J.M. Yeh, L. Diller, Pediatric Hodgkin lymphoma: trade-offs between short- and long-term mortality risks, Blood 120 (2012) 2195–2202, https://doi.org/10.1182/blood-2012-02-409821

[11] R.A. Kleinerman, Cancer risks following diagnostic and therapeutic radiation exposure in children, Pediatr. Radiol. 36 (2006) 121–125, https://doi.org/10.1007/s00247-006-0191-5

[12] J. Stauss, C. Franzius, T. Pfluger, K.U. Juergens, L. Biassoni, J. Begent, R. Kluge, H. Amthauer, T. Voelker, L. H??jgaard, S. Barrington, S. Hain, T. Lynch, K. Hahn, Guidelines for 18F-FDG PET and PET-CT imaging in paediatric oncology, Eur. J. Nucl. Med. Mol. Imaging 35 (2008), pp. 1581–1588, https://doi.org/10.1007/s00259-008-0826-x

[13] D.L. Miglioretti, E. Johnson, A. Williams, T. Robert, S. Weinmann, L.I. Solberg, H. Spencer, D. Roblin, M.J. Flynn, N. Vanneman, Pediatric computed tomography and associated radiation exposure and estimated cancer risk, JAMA Pediatr. 167 (2017) 6–14, https://doi.org/10.1001/jamapediatrics.2013.311.Pediatric

[14] R. Smith-Bindman, J. Lipson, R. Marcus, K.P. Kim, M. Mahesh, R. Gould, A.B. de Gonzalez, D.L. Miglioretti, Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer, Arch Intern Med. 118 (2016) 6072–6078, https://doi.org/10.1002/cncr.27633.Percutaneous

[15] A. Sud, K. Hemminki, R.S. Houlston, Second cancer risk following Hodgkin lymphoma, Oncotarget 8 (2017) 78261–78262, https://doi.org/10.18632/oncotarget.20876

[16] D.J. Brenner, C.D. Elliston, E.J. Hall, W.E. Berdon, Estimated risks of radiation-induced fatal cancer from pediatric CT, Am. Roentgen Ray Soc. 176 (2001) 289–296, https://doi.org/10.2214/ajr.176.2.1760289

[17] B. Huang, M.W.M. Law, P.L. Khong, Whole-body PET/CT scanning: estimation of radiation dose and cancer risk, Radiology 251 (2009) 166–174, https://doi.org/10.1148/radiol.2511081300

[18] S.-H. Chien, C.-J. Liu, Y.-W. Hu, Y.-C. Hong, C.-J. Teng, C.-M. Yeh, T.-J. Chiou, J.-P. Gau, C.-H. Tzeng, Frequency of surveillance computed tomography in non-Hodgkin lymphoma and the risk, Int. J. Cancer. (2015) 658–665.

[19] P. Shenoy, R. Sinha, J.W. Tumeh, M.J. Lechowicz, C. Flowers, Surveillance computed tomography scans for patients with lymphoma: Is the risk worth the benefits? Clin. Lymphoma, Myeloma Leuk 10 (2010) 270–277, https://doi.org/10.3816/CLML.2010.n.056

[20] D.J. Brenner, E.J. Hall, Computed tomography - an increasing source of radiation exposure, N. Engl. J. Med. (2007) 2277–2284.

[21] A.S. Littooij, T.C. Kwee, I. Barber, C. Granata, M.A. Vermoolen, G. Enríquez, J. Zsíros, S.Y. Soh, B. De Keizer, F.J.A. Beek, M.G. Hobbelink, M.B. Bierings, J. Stoker, R.A.J. Nievelstein, Whole-body MRI for initial staging of paediatric lymphoma: prospective comparison to an FDG-PET/CT-based reference standard, Eur. Radiol. 24 (2014) 1153–1165, https://doi.org/10.1007/s00330-014-3114-0

[22] A. Latifoltojar, S. Punwani, A. Lopes, P.D. Humphries, M. Klusmann, L.J. Menezes, S. Daw, A. Shankar, D. Neriman, H. Fitzke, L. Clifton-Hadley, P. Smith, S.A. Taylor, Whole-body MRI for staging and interim response monitoring in paediatric and adolescent Hodgkin's lymphoma: a comparison with multi-modality reference standard including18F-FDG-PET-CT, Eur. Radiol. (2019) 1–11, https://doi.org/10.1007/s00330-018-5445-8

[23] D. Albano, C. Patti, L. La Grutta, F. Agnello, E. Grassedonio, A. Mulè, G. Cannizzaro, U. Ficola, R. Lagalla, M. Midiri, M. Galia, Comparison between whole-body MRI with diffusion-weighted imaging and PET/CT in staging newly diagnosed FDG-avid lymphomas. Eur. J. Radiol. 85 (2016) 313–318, https://doi.org/10.1016/j.ejrad.2015.12.006

[24] A. Balbo-Mussetto, S. Cirillo, R. Bruna, A. Gueli, C. Saviolo, M. Petracchini, A. Fornari, C.V. Lario, D. Gottardi, A. De Crescenzo, C. Tarella, Whole-body MRI with diffusion-weighted imaging: a valuable alternative to contrast-enhanced CT for initial staging of aggressive lymphoma, Clin. Radiol. 71 (2016) 271–279, https://doi.org/10.1016/j.crad.2015.11.018

[25] C. Ferrari, C. Minoia, A.N. Asabella, A. Nicoletti, C. Altini, F. Antonica, M. Ficco, A. Guarini, N. Maggialetti, G. Rubini, 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. 17 (Suppl 1) (2014) 40–49 〈http://www.ncbi.nlm.nih.gov/pubmed/24392468〉 accessed November 14, 2017.

[26] M.E. Mayerhoefer, G. Karanikas, K. Kletter, H. Prosch, B. Kiesewetter, C. Skrabs, E. Porpaczy, M. Weber, K. Pinker-Domenig, D. Berzaczy, M. Hoffmann, C. Sillaber, U. Jaeger, L. Mullauer, I. Simonitsch-Klupp, W. Dolak, A. Gaiger, P. Ubl, J. Lukas, M. Raderer, Evaluation of diffusion-weighted MRI for pretherapeutic assessment and staging of lymphoma: results of a prospective study in 140 patients, Clin. Cancer Res 20 (2014) 2984–2993, https://doi.org/10.1158/1078-0432.CCR-13-3355

[27] S. Spijkers, A.S. Littooij, T.C. Kwee, N. Tolboom, A. Beishuizen, M.C.A. Bruin, S.G. Elias, T. van de Brug, G. Enriquez, C. Sábado, E. Miller, C. Granata, C. de Lange, F. Verzegnassi, M.-L.C. Greer, B. de Keizer, R.A.J. Nievelstein, Whole-body MRI versus an FDG-PET/CT-based reference standard for staging of paediatric Hodgkin lymphoma: a prospective multicentre study, Eur. Radiol (2020), https://doi.org/10.1007/s00330-020-07182-0

[28] S. Punwani, S.A. Taylor, Z.Z. Saad, A. Bainbridge, A. Groves, S. Daw, A. Shankar, S. Halligan, P.D. Humphries, Diffusion-weighted MRI of lymphoma: prognostic utility and implications for PET/MRI? Eur. J. Nucl. Med. Mol. Imaging 40 (2013) 373–385, https://doi.org/10.1007/s00259-012-2293-7

[29] R. Regacini, A. Puchnick, D.C. Shigueoka, W. Iared, H.M. Lederman, Whole-body diffusion-weighted magnetic resonance imaging versus FDG-PET/CT for initial lymphoma staging: systematic review on diagnostic test accuracy studies, Sao Paulo Med. J 133 (2015) 141–150, https://doi.org/10.1590/1516-3180.2014.8312810

[30] D. Wang, Y. Huo, S. Chen, H. Wang, Y. Ding, X. Zhu, C. Ma, Whole-body MRI versus 18F-FDG PET/CT for pretherapeutic assessment and staging of lymphoma: a meta-analysis, Oncol. Targets. Ther Volume 11 (2018) 3597–3608, https://doi.org/10.2147/ott.s148189

[31] P.F. Whiting, A.W.S. Rutjes, M.E. Westwood, S. Mallett, J.J. Deeks, J.B. Reitsma, M.M.G. Leeflang, J.A.C. Sterne, P.M.M. Bossuyt, QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies, Ann. Intern. Med. 155 (2011) 529–536.

[32] E.B. Wilson, Probable inference, the law of succession, and statistical inference, J. Am. Stat. Assoc. 22 (2016) 209–212.

[33] J.P.T. Higgins, S.G. Thompson, J.J. Deeks, D.G. Altman, Measuring inconsistency in knowledgebases, J. Intell. Inf. Syst 27 (2006) 159–184, https://doi.org/10.1007/s10844-006-2974-4

[34] S. Punwani, S.A. Taylor, A. Bainbridge, V. Prakash, S. Bandula, E. De Vita, O.E. Olsen, S.F. Hain, N. Stevens, S. Daw, A. Shankar, J.B. Bomanji, P.D. Humphries, Pediatric and adolescent lymphoma: comparison of whole-body STIR half-fourier RARE MR imaging with an enhanced PET/CT reference for initial staging, Radiology 255 (2010) 182–190, https://doi.org/10.1148/radiol.09091105

[35] S. Kharuzhyk, E. Zhavrid, A. Dziuban, E. Sukolinskaja, O. Kalenik, Comparison of whole-body MRI with diffusion-weighted imaging and PET/CT in lymphoma staging, Eur. Radiol. (2020), https://doi.org/10.1007/s00330-020-06732-w

[36] K. Herrmann, M. Queiroz, M.W. Huellner, F. de Galiza Barbosa, A. Buck, N. Schaefer, P. Stolzman, P. Veit-Haibach, Diagnostic performance of FDG-PET/MRI and WB-DW-MRI in the evaluation of lymphoma: a prospective comparison to standard FDG-PET/CT, BMC Cancer 15 (2015) 1002, https://doi.org/10.1186/s12885-015-2009-z

[37] A.S. Littooij, T.C. Kwee, B. De Keizer, M.C.A. Bruin, A. Coma, F.J.A. Beek, R. Fijnheer, R.A.J. Nievelstein, Whole-body MRI-DWI for assessment of residual disease after completion of therapy in lymphoma: A prospective multicenter study, J. Magn. Reson. Imaging 42 (2015) 1646–1655, https://doi.org/10.1002/jmri.24938

[38] T. Hagtvedt, T. Seierstad, K.V. Lund, A.M. Løndalen, T.V. Bogsrud, H.J. Smith, O.M. Geier, H. Holte, T.M. Aaløkken, Diffusion-weighted MRI compared to FDG PET/CT for assessment of early treatment response in lymphoma, Acta Radiol 56 (2014) 152–158, https://doi.org/10.1177/0284185114526087

[39] M.E. Mayerhoefer, G. Karanikas, K. Kletter, H. Prosch, B. Kiesewetter, C. Skrabs, E. Porpaczy, M. Weber, T. Knogler, C. Sillaber, U. Jaeger, I. Simonitsch-Klupp, P. Ubl, L. Müllauer, W. Dolak, J. Lukas, M. Raderer, Evaluation of diffusion-weighted magnetic resonance imaging for follow-up and treatment response assessment of lymphoma: Results of an 18F-FDG-PET/CT-controlled prospective study in 64 patients, Clin. Cancer Res 21 (2015) 2506–2513, https://doi.org/10.1158/1078-0432.CCR-14-2454

[40] N. Maggialetti, C. Ferrari, C. Minoia, A.N. Asabella, M. Ficco, G. Loseto, G. De Tullio, V. de Fazio, A. Calabrese, A. Guarini, G. Rubini, L. Brunese, Role of WB-MR/DWIBS compared to 18F-FDG PET/CT in the therapy response assessment of lymphoma, Radiol. Medica. 121 (2016) 132–143, https://doi.org/10.1007/s11547-015-0581-6

[41] A.J. Theruvath, F. Siedek, A.M. Muehe, J. Garcia-diaz, J. Kirchner, O. Martin, M.P. Link, S. Spunt, J.F. Schäfer, M. Moseley, L. Umutly, H.E. Daldrup-Link, Therapy Response Assessment of Pediatric Tumors with Whole-Body Diffusion-weighted MRI and FDG PET/MRI, Radiology (2020) 1–9, https://doi.org/10.1148/radiol.2020192508

[42] S. Spijkers, A.S. Littooij, T.C. Kwee, N. Tolboom, A. Beishuizen, M.C.A. Bruin, G. Enriquez, C. Sabado, E. Miller, C. Granata, C. de Lange, F. Verzegnassi, B. de Keizer, R.A.J. Nievelstein, Whole-body MRI versus an [18F]FDG-PET/CT-based reference standard for early response assessment and restaging of paediatric Hodgkin's lymphoma: a prospective multicentre study, Dec 31, Eur Radiol 12 (2021) 8925–8936, https://doi.org/10.1007/s00330-021-08026-1

[43] J.F. Schäfer, C. Granata, T. von Kalle, M. Kyncl, A.S. Littooij, P.L. Di Paolo, I. Sefic Pasic, R.A.J. Nievelstein, Oncology Task Force of the ESPR. Whole-body magnetic resonance imaging in pediatric oncology - recommendations by the Oncology Task Force of the ESPR (Jul), Pediatr Radiol 50 (8) (2020) 1162–1174, https://doi.org/10.1007/s00247-020-04683-4

[44] A.S. Littooij, T.C. Kwee, I. Barber, C. Granata, B. de Keizer, F.J. Beek, M.G. Hobbelink, R. Fijnheer, J. Stoker, R.A. Nievelstein, Accuracy of whole-body MRI in the assessment of splenic involvement in lymphoma, Acta Radiol 57 (2015) 1–10, https://doi.org/10.1177/0284185115571657

[45] L.M. Wu, F.Y. Chen, X.X. Jiang, H.Y. Gu, Y. Yin, J.R. Xu, 18F-FDG PET, combined FDG-PET/CT and MRI for evaluation of bone marrow infiltration in staging of

lymphoma: A systematic review and meta-analysis, Eur. J. Radiol 81 (2012) 303–311, https://doi.org/10.1016/j.ejrad.2010.11.020

[46] D. Albano, C. Patti, R. Lagalla, M. Midiri, M. Galia, Whole-body MRI, FDG-PET/CT, and bone marrow biopsy, for the assessment of bone marrow involvement in patients with newly diagnosed lymphoma, J. Magn. Reson. Imaging 45 (2017) 1082–1089, https://doi.org/10.1002/jmri.25439

[47] E. Even-Sapir, G. Lievshitz, C. Perry, Y. Herishanu, H. Lerman, U. Metser, Fluorine-18 Fluorodeoxyglucose PET/CT Patterns of Extranodal Involvement in Patients with Non-Hodgkin Lymphoma and Hodgkin's Disease, Radiol. Clin. North Am. 45 (2007) 697–709, https://doi.org/10.1016/j.rcl.2007.05.009

[48] A. Balbo-Mussetto, C. Saviolo, A. Fornari, D. Gottardi, M. Petracchini, A. Macera, C.V. Lario, T. Gallo, C. Tarella, S. Cirillo, Whole body MRI with qualitative and quantitative analysis of DWI for assessment of bone marrow involvement in lymphoma, Radiol. Medica. 122 (2017) 623–632, https://doi.org/10.1007/s11547-017-0762-6