

Whole-body diffusion-weighted magnetic resonance imaging versus FDG-PET/CT for initial lymphoma staging: systematic review on diagnostic test accuracy studies

Ressonância magnética de corpo total com difusão comparada à FDG-PET/CT no estadiamento inicial do linfoma: revisão sistemática de estudos de acurácia diagnóstica

Rodrigo Regacini^I, Andrea Puchnick^{II}, David Carlos Shigueoka^{III}, Wagner Iared^{IV}, Henrique Manoel Lederman^V

Department of Diagnostic Imaging, Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil

^IMD, MSc. Radiologist, Discipline of Pediatric Radiology, Department of Diagnostic Imaging, Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil.

^{II}BSc. Professor and Coordinator of Educational and Research Support, Department of Diagnostic Imaging, Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil.

^{III}MD, PhD. Adjunct Professor, Department of Diagnostic Imaging, Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil.

^{IV}MD, PhD. Assistant Research Radiologist, Department of Diagnostic Imaging, Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil.

^VMD, PhD. Full Professor and Head of the Discipline of Pediatric Radiology, Department of Diagnostic Imaging, Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil.

KEY WORDS:

Magnetic resonance imaging.
Diffusion magnetic resonance imaging.
Whole body imaging.
Lymphoma.
Positron-emission tomography.

PALAVRAS-CHAVE:

Imagem por ressonância magnética.
Imagem de difusão por ressonância magnética.
Imagem corporal total.
Linfoma.
Tomografia por emissão de pósitrons.

ABSTRACT

CONTEXT AND OBJECTIVE: Positron emission tomography with [18F]-fluoro-2-deoxyglucose (FDG-PET/CT) has been advocated as the method of choice for lymphoma staging, since it enables whole-body analysis with high sensitivity for detection of affected areas and because it combines capacities for anatomical and functional assessment. With technological advances, magnetic resonance imaging (MRI) has emerged as an alternative to FDG-PET/CT. This systematic review with meta-analysis aimed to compare whole-body diffusion-weighted MRI (WB-MRI) with FDG-PET/CT for lymphoma staging.

DESIGN AND SETTING: Systematic review on diagnostic test accuracy studies conducted at a public university.

METHODS: The Medline, Scopus, Embase and Lilacs databases were searched for studies published up to September 2013 that compared WB-MRI and FDG-PET/CT for lymphoma staging. The reference lists of included studies were checked for any relevant additional citations.

RESULTS: Six studies that evaluated the initial lymphoma staging in 116 patients were included. WB-MRI and FDG-PET/CT agreed in 90.5% of the cases ($\kappa = 0.871$; $P < 0.0001$). In most of the studies, when there was disagreement between the methods, WB-MRI overstaged in relation to FDG-PET/CT. The sensitivity of WB-MRI and FDG-PET/CT, in comparison with the clinical-radiological standard, ranged from 59 to 100% and from 63 to 100% respectively.

CONCLUSION: WB-MRI is a highly sensitive method for initial lymphoma staging. It has excellent agreement with FDG-PET/CT and is a great alternative for managing lymphoma patients, without using ionizing radiation or an intravenous contrast agent.

RESUMO

CONTEXTO E OBJETIVO: A tomografia por emissão de pósitrons com 2-[18F]-fluoro-2-deoxi-D-glicose (FDG-PET/CT) tem sido defendida como método de escolha para o estadiamento do linfoma por realizar o estudo do corpo inteiro com boa sensibilidade para detecção das áreas acometidas e por combinar as capacidades de avaliação anatômica e funcional. Com os avanços tecnológicos, a ressonância magnética tem se apresentando como alternativa à FDG-PET/CT. Esta revisão sistemática com metanálise visa comparar a ressonância magnética de corpo inteiro (WB-MRI) com difusão com a FDG-PET/CT no estadiamento do linfoma.

TIPO DE ESTUDO E LOCAL: Revisão sistemática de estudos de acurácia diagnóstica conduzida em universidade pública.

MÉTODOS: Foi conduzida uma busca nos bancos de dados Medline, Embase, Scopus e Lilacs por estudos publicados até setembro de 2013 comparando a WB-MRI com a FDG-PET/CT no estadiamento do linfoma. As referências bibliográficas dos estudos incluídos foram checadas com a finalidade de encontrar citações adicionais relevantes.

RESULTADOS: Foram incluídos seis estudos que avaliaram o estadiamento inicial do linfoma de 116 pacientes. A WB-MRI e a FDG-PET/CT concordaram em 90,5% dos casos ($\kappa = 0,871$; $P < 0,0001$). Na maioria dos estudos, quando houve discordância, a WB-MRI estabeleceu estadiamento superior à FDG-PET/CT. A sensibilidade da WB-MRI e da FDG-PET/CT, em relação ao padrão clínico-radiológico, variou de 59% a 100% e de 63% a 100%, respectivamente.

CONCLUSÃO: A WB-MRI apresenta alta sensibilidade no estadiamento inicial do linfoma, excelente concordância com a FDG-PET/CT e representa uma ótima alternativa no manejo de pacientes com linfoma, sem utilizar radiação ionizante ou meio de contraste intravenoso.

INTRODUCTION

Lymphomas account for approximately 5-6% of all malignancies.¹ Over two-thirds of these cases are non-Hodgkin lymphomas (NHL), and Hodgkin's lymphoma (HL) makes up the rest.¹ After a histopathological diagnosis has been established, the imaging-based initial staging will influence the choice of therapy and prognosis, aid in radiation therapy planning for localized disease and provide a baseline for treatment response monitoring.^{2,3} HL and NHL staging is currently based on the Cotswolds modification of the Ann Arbor classification system.⁴ This system uses the number of tumor sites, the extent of involvement (nodal or extranodal) and its distribution as staging factors, whereas the Cotswolds modification also takes tumor burden into account.

Several imaging methods have been used for this purpose and, of these, computed tomography (CT) is currently the most popular.^{2,3} Over recent years, [18]F-fluoro-2-deoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) has emerged as the most accurate method of all. It is based on the principle that malignant tissues exhibit higher glucose metabolism than that of healthy tissue⁵ and enables whole-body scanning with high sensitivity for detection of affected areas while combining the anatomical and functional assessment capabilities of CT and PET.^{6,7} However, its sensitivity and specificity vary according to histological subtype,^{8,9} and use of PET/CT has been correlated with substantial radiation exposure, particularly because scans must often be obtained repeatedly over the treatment course. Recent studies have shown that radiation exposure secondary to diagnostic imaging leads to increased lifetime risk of malignant tumors, especially in children.¹⁰⁻¹²

Magnetic resonance imaging (MRI) has emerged as a safer alternative for lymphoma staging, since progress in MRI techniques now enables rapid whole-body scanning¹³ while potentially providing the same information as FDG-PET/CT.^{14,15} The functional assessment in whole-body MRI (WB-MRI) is based on diffusion-weighted imaging (DWI), a method that maps water molecule movement in tissue (within cells, in the extracellular medium and across cell membranes). In the presence of lymphomas, the Brownian motion of water molecules is restricted due to increased tissue cellularity and elevated nucleus-to-cytoplasm ratio, which will produce relatively high signal intensity on DWI, compared with normal tissues.¹⁶ Using this principle, diffusion MRI can detect tumor-related changes that are not limited to anatomical information.¹⁷ Furthermore, apparent diffusion coefficient (ADC) quantification on DWI can provide useful information on treatment response and help distinguish benign from malignant lymph nodes.¹⁸

Over the last decade, a growing number of studies have compared WB-MRI and FDG-PET/CT in patients with

lymphoma, using a variety of approaches. In studies focusing solely on initial lymphoma staging, the two methods are usually compared in two ways: taking into account the accuracy of each method for detection of individual lesions (on the basis of the number of lesions detected); or taking into account the final staging score, regardless of the number of lesions detected through each method.

Comparative analysis on these studies can be quite challenging when this attempts to focus on the ability of each method to detect individual lesions. The difficulty is mostly due to the wide range of WB-MRI protocols used, which precludes proper comparison. However, since the ultimate objective of initial lymphoma imaging is to define the disease stage at baseline, studies can be compared on the basis of the staging scores indicated by each method, regardless of the number of lesions detected.

OBJECTIVES

Within this context, this study aimed to compare whole-body diffusion-weighted MRI (WB-MRI) with PET/CT for lymphoma staging by means of a meta-analysis, in order to identify whether the data available in the literature are sufficient to establish that WB-MRI is a safe alternative for lymphoma staging.

METHODS

Type of study and participants

This was a systematic review of diagnostic test accuracy studies, with meta-analysis. The spectrum of patients included HL and NHL cases.

The present study was approved by the local Research Ethics Committee, under number 0135/12HE.

Inclusion criteria

All diagnostic test accuracy studies, comparing WB-MRI versus FDG-PET/CT for initial lymphoma staging, with the added utility of DWI in WB-MRI, which were published up to September 2013, were assessed.

Exclusion criteria

Studies meeting any of the following criteria were excluded: data could only be extracted for one of the methods under analysis; FDG-PET/CT was used as the single reference standard for lymphoma staging; samples included cases previously reported elsewhere; the data represented a subpopulation analysis from larger investigations previously included in our review; the study included patients with diseases other than lymphoma; or the study assessed the performance of WB-MRI in relation to lymphomas, but only for detection of bone involvement.

Search strategy

The Medline (via PubMed), Embase, Lilacs and Scopus databases were searched for relevant studies on the performance of WB-MRI versus other imaging methods for lymphoma evaluation. The references of each study included were checked for potentially relevant additional citations. The results from our search strategy are shown in **Table 1**. The search was last updated on September 27, 2013.

Article selection and quality assessment

For the first stage of the selection, two investigators (RR, AP) conducted independent assessments of the titles and abstracts of articles identified by the abovementioned search strategy. Studies on the diagnostic performance of WB-MRI for lymphoma staging or follow-up were included. Animal studies, reviews, meta-analyses, abstracts, editorials, letters to the editor, case reports, tutorials and practice guidelines were excluded. All clearly ineligible articles were also excluded.

For the second stage, all potentially eligible studies were set aside for full-text reading, critical appraisal and data extraction, conducted independently by the same investigators (RR, AP). Any disagreements arising between them at either stage were resolved through discussion and reaching a consensus.

Study quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2).¹⁹ The QUADAS-2 tool enables more transparent ratings for bias and for the applicability of diagnostic accuracy studies. Three responses to questions regarding the risk of bias and applicability concerns were possible: “low”, “high” or “unclear”.

Reference standard

Lymphoma staging provided by a set of clinical and radiological data was used to compare each technique in accordance with the Ann Arbor staging system. The final staging established needed to take into account all the clinical information available at the time of diagnosis, such as physical examination, laboratory and histological results, and bone marrow biopsy, and also the information available during the clinical and imaging follow-up (CT, FDG-PET/CT, WB-MRI or other methods). This follow-up was used to determine the status of lesions after treatment. For example, if they became larger during the follow-up period or

decreased in size after treatment, they were considered positive for the presence of lymphoma.

Statistical analysis

The Review Manager (RevMan) version 5.1 (Cochrane Collaboration, Oxford, England) was used to calculate sensitivity and specificity, with 95% confidence intervals (CIs), for WB-MRI and FDG-PET/CT, in comparison with a clinical-radiological standard reference. The results from each individual study were presented in forest plots.

Statistical analyses were performed using the SPSS 16.0 software package. The unweighted kappa (κ) statistic was used to test agreement between WB-MRI and FDG-PET/CT in the initial lymphoma staging. Agreement was considered poor at a κ value of 0, weak at 0.01-0.20, fair at 0.21-0.40, moderate at 0.41-0.60, good at 0.61-0.80 and excellent at 0.81-1.0.²⁰ P values < 0.05 were considered indicative of significant differences.

RESULTS

The search strategy chosen yielded 929 citations. After careful reading of the titles and abstracts and exclusion of duplicates, 19 articles were selected for full-text analysis and critical appraisal. Thirteen failed to meet the inclusion criteria (or met the exclusion criteria) and were excluded from further analysis: one study compared WB-MRI versus FDG-PET/CT for assessment of treatment response across a patient spectrum previously used in another study included in this systematic review;²¹ one study used FDG-PET/CT as a single reference standard for lymphoma staging;²² two studies assessed the performance of WB-MRI in lymphoma, but only for detection of bone involvement;^{23,24} two studies only compared WB-MRI with conventional CT;^{25,26} four studies had samples that included patients with diseases other than lymphoma;²⁷⁻³⁰ one study only compared WB-MRI with a reference standard;³¹ and two studies only compared WB-MRI with conventional CT and bone scintigraphy.^{32,33}

On completion of the search and retrieval strategy, six prospective cohort studies were included for meta-analysis.^{14-16,34-36} Most of them either failed to conduct separate analyses on HL and NHL or conducted pooled analyses on different histological subtypes of NHL. The study by Wu et al.³⁴ limited its analysis only to a single histological subtype of NHL. **Table 2** provides a summary of the key features of these studies.

Table 1. Search strategy results

| # | Search query | PubMed | Embase | Scopus | Lilacs |
|---|--|---------|---------|---------|--------|
| 1 | “Magnetic Resonance Imaging” OR “Diffusion Magnetic Resonance Imaging” | 365,239 | 524,605 | 689,460 | 2,348 |
| 2 | “Whole Body Imaging” | 72,832 | 93,962 | 145,328 | 25 |
| 3 | “Lymphoma” OR “Hodgkin Disease” OR “Lymphoma, Non-Hodgkin” | 210,131 | 215,548 | 456,017 | 1,551 |
| 4 | (#1 AND #2) AND #3 | 184 | 355 | 390 | 0 |

Note: all synonyms under the hierarchical tree of each term were also included for each database.

Quality assessment on studies included

Figure 1 summarizes the risk of bias and applicability judgments on the six studies included. The methodological quality graph presents the percentage of included studies for which the item was rated “low”, “high” or “unclear”, for each quality assessment domain. The graph shows that the potential area of concern was the description of the reference standard.

Patient selection criteria were clearly described in all the studies included. Regarding the reference standard, van Ufford et al.³⁵ and Lin et al.³⁶ did not describe it clearly. All the other quality assessment parameters were considered satisfactory across all six studies.

The QUADAS-2 score, expressed as a percentage of the maximum score, was 90% on average (range, 71-100%) in the six studies included. In the quality assessment, all of the studies were considered to present low risk of bias and low concerns about applicability.

Summary assessment of the sensitivity of WB-MRI for lymphoma staging

The sensitivity of WB-MRI and FDG-PET/CT for initial lymphoma staging versus that of the reference standard ranged from 59% to 100% and from 63% to 100% respectively (Figure 2). Gu et al.,¹⁴ Abdulqadhr et al.,¹⁵ Stéphane et al.¹⁶ and Lin et al.³⁶

Table 2. Key features of the studies included in this meta-analysis (all of them were prospective cohort studies)

| Study | Participants | Target condition | Dates (country) |
|---------------------------------|--|----------------------------------|--|
| Gu et al. ¹⁴ | 17 patients (11 M, 6 F); mean age, 50 years (SD, 17 years; range, 20-80 years) | HL and NHL | November 2008-April 2010 (China) |
| Abdulqadhr et al. ¹⁵ | 31 patients (20 M, 11 F); mean age, 47.2 years (range, 18-78 years) | HL, aggressive NHL, indolent NHL | March 2008-November 2009 (Sweden) |
| Stéphane et al. ¹⁶ | 23 patients (11 M, 12 F); mean age, 51 years (range, 18-84 years) | HL, DLBCL and NHL | June 2008-October 2009 (France) |
| Wu et al. ³⁴ | 8 patients (4 M, 4 F); age range, 32-78 years | DLBCL | N/D (Finland) |
| van Ufford et al. ³⁵ | 22 patients (16 M, 6 F); mean age, 59.9 years (SD, 14.3 years; range, 22-81 years) | HL and NHL | August 2008-October 2009 (Netherlands) |
| Lin et al. ³⁶ | 15 patients (9 M, 6 F); mean age, 48 years (range, 23-79 years) | DLBCL | June 2008-February 2009 (France) |

M = male; F = female; SD = standard deviation; HL = Hodgkin's lymphoma; NHL = non-Hodgkin lymphoma; DLBCL = diffuse large B-cell lymphoma; N/D = no data.

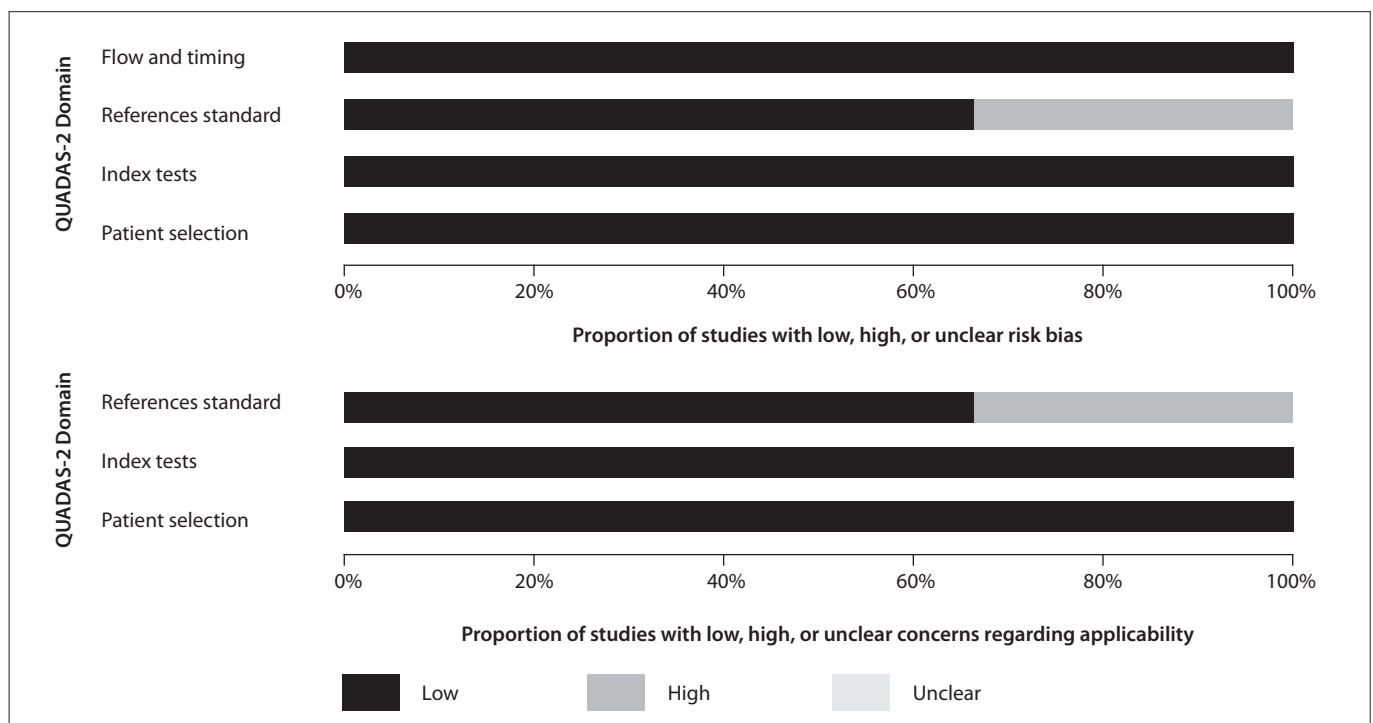


Figure 1. Graphical representation of study quality assessment.

reported high sensitivity for both methods, whereas Wu et al.³⁴ and van Ufford et al.³⁵ found lower sensitivity values.

Agreement between WB-MRI and FDG-PET/CT for lymphoma staging

In the study by Gu et al.,¹⁴ there was agreement between WB-MRI and FDG-PET/CT staging in 15 of their 17 patients. In the remaining two patients, WB-MRI overstaged one and understaged the other. In the latter patient, the staging with both methods was considered inadequate in relation to the reference standard because both of them failed to detect bone marrow infiltration, which was later confirmed by means of bone marrow biopsy. In the study by Abdulqadhr et al.,¹⁵ there was agreement between WB-MRI and FDG-PET/CT staging in 28 of their 31 patients. In the remaining three patients, low-grade lymphoma had higher staging through WB-MRI than through FDG-PET/CT, which was later validated by means of clinical staging. In the studies by Stéphane et al.¹⁶ and Wu et al.³⁴ WB-MRI and FDG-PET/CT yielded the same staging in all patients, although three were incorrectly staged with both methods in the study by Wu et al.³⁴ In the sample of van Ufford et al.,³⁵ WB-MRI and FDG-PET/CT agreed regarding the staging of 17 of their 22 patients. WB-MRI overstaged five patients in relation to FDG-PET/CT, and only one of these patients, who had bone marrow infiltration later confirmed by biopsy, was correctly staged by means of the imaging method. Finally, in the study by Lin et al.,³⁶ WB-MRI and FDG-PET/CT yielded similar staging for 14 patients. In the sole case in which the staging was different between the methods, it was higher with WB-MRI than with FDG-PET/CT, although

both methods staged the patient incorrectly, compared with the reference standard. The kappa statistic was indicative of excellent overall agreement between WB-MRI and FDG-PET/CT ($\kappa = 0.871$ [0.782; 0.960]; $P < 0.0001$). Table 3 summarizes the agreement between the two methods in each study.

DISCUSSION

The studies compared WB-MRI and FDG-PET/CT in terms of their ability to detect sites of disease involvement (both nodal and extranodal). This type of analysis is based on detection of individual lesions, i.e. each lesion regarded as positive counts toward the analysis on the agreement between the methods. Lesion-by-lesion comparison hinders meta-analysis on these studies, particularly because different methods are used in different studies. Differences in WB-MRI protocols, which may use distinct sequences with different acquisition planes, slice

Table 3. Agreement between WB-MRI and FDG-PET/CT for lymphoma staging in each of the studies included

| Study | Kappa coefficient | | | P |
|---------------------------------|-------------------|----------------|-----------|----------|
| | κ | CI | Agreement | |
| Gu et al. ¹⁴ | 0.824 | [0.707; 0.941] | Excellent | < 0.0001 |
| Abdulqadhr et al. ¹⁵ | 0.869 | [0.796; 0.941] | Excellent | < 0.0001 |
| Stéphane et al. ¹⁶ | 1.000 | — | Excellent | < 0.0001 |
| Wu et al. ³⁴ | 1.000 | — | Excellent | < 0.0001 |
| van Ufford et al. ³⁵ | 0.683 | [0.564; 0.803] | Good | < 0.0001 |
| Lin et al. ³⁶ | 0.847 | [0.712; 0.981] | Excellent | < 0.0001 |

WB-MRI = whole-body diffusion-weighted magnetic resonance imaging; FDG-PET/CT = [18]F-fluoro-2-deoxyglucose positron emission tomography; CI = confidence interval.

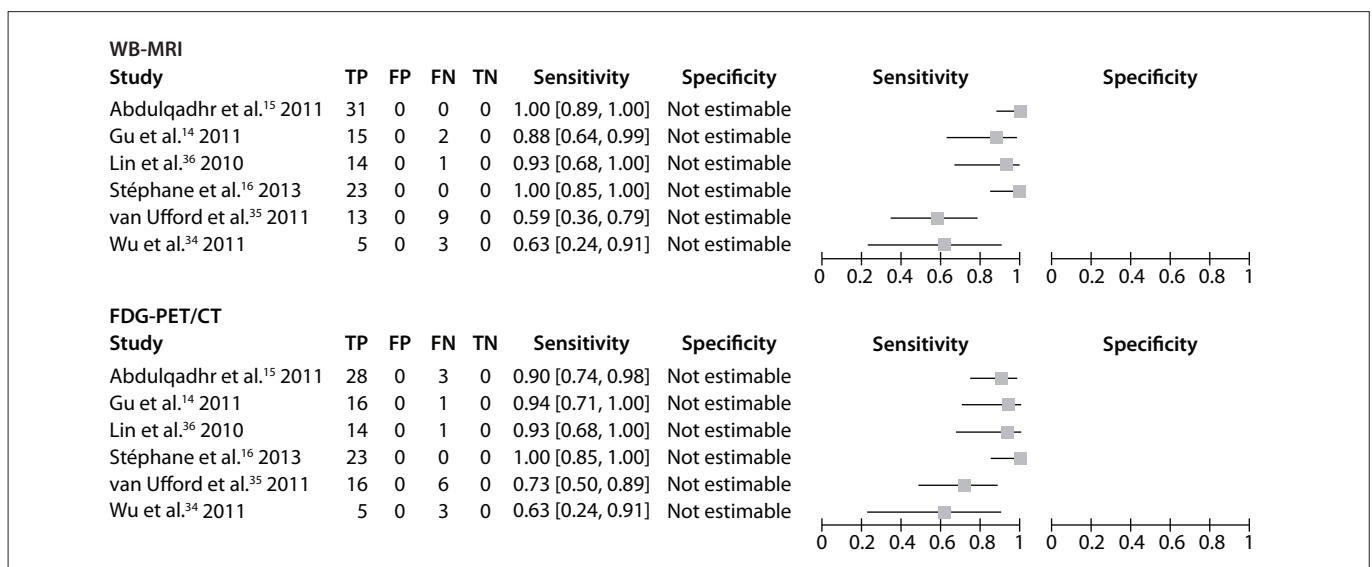


Figure 2. Forest plots of the sensitivity of imaging methods for lymphoma staging versus a comparison reference standard. TP = true positive, FP = false positive, FN = false negative, TN = true negative. Brackets show 95% confidence intervals. The figure shows the sensitivity for each study (squares) and 95% confidence intervals (horizontal lines). Specificity was not calculable, since all patients had lymphoma.

thicknesses and body areas (Table 4) are a particular cause for concern. Lymph node size cutoffs and the criteria used to classify an organ or extranodal lesion as “involved” also differed across studies. In view of this heterogeneity, lesions that were considered positive with one imaging method may have been classified as negative with the other.

However, all the studies also conducted staging in accordance with the Ann Arbor classification, which is based on a clinical and radiological reference standard, to which both WB-MRI and FDG-PET/CT staging were compared. Lymphoma staging in the

Ann Arbor system is dependent on disease distribution (above or below the diaphragm) and on the affected lymph node sites. Detection of one abnormal lymph node or extranodal lesion suffices for a region to be classified as involved. Hence, for staging purposes, detection of additional lesions in a region or organ classified as positive is of no utility. This enables comparison between studies, despite methodological differences, in terms of the ability of each imaging method to establish staging.

Overall, the six studies included in this meta-analysis assessed 116 patients. There was agreement in staging between

Table 4. WB-MRI protocols used in studies included in the meta-analysis

| Study | Scanner | | Sequence | Protocol | | Extent |
|---------------------------------|--|---|----------|---|----------------------------------|--|
| | Field strength | Coils used | | Parameters | Plane | |
| Gu et al. ¹⁴ | 3 T | Surface: abdomen, pelvis Body: head, neck, chest | T2 | Slice thickness: 5 mm | Axial | Head to distal thigh |
| | | | T2 SPAIR | Slice thickness: 5 mm | | |
| | | | DWI | Slice thickness: 5 mm b-value: 0 and 1000 s/mm ² | | |
| Abdulqadhr et al. ¹⁵ | 1.5 T | Built-in body coil | T1 | In-phase and out-phase Slice thickness: 6 mm | Coronal | Not reported |
| | | | T2 STIR | Slice thickness: 6 mm | | |
| | | | DWI | Slice thickness: 6 mm b-value: 0 and 1000 s/mm ² ADC not calculated | Axial with 3D MIP reconstruction | |
| Stéphane et al. ¹⁶ | 1.5 T | Built-in body coil | T1 | Not reported | Not reported | Not reported |
| | | | STIR | Not reported | Not reported | Not reported |
| | | | DWI | Slice thickness: 4 mm b-value: 0 and 1000 s/mm ² ADC calculated in the axillary and inguinal regions | Axial with 3D MIP reconstruction | From the skull base to the thigh |
| Wu et al. ³⁴ | 3 T | Body and spine | T1 | Slice thickness: 3 mm | Coronal | Neck, chest, abdomen, pelvis |
| | | | T2 IR | Slice thickness: 3 mm | Coronal | |
| | | | T1 | Slice thickness: 1 mm | Axial | Neck |
| | | | T2 | Slice thickness: 3 mm | | |
| | | | T2 SAT | Slice thickness: 3 mm Slice thickness: 5 mm | | |
| | | | DWI | b-value: 0 and 800 s/mm ² ADC calculated | | |
| | | | T1 | Slice thickness: 2 mm | | |
| T2 | Slice thickness: 3 mm | | | | | |
| T2 SAT | Slice thickness: 3 mm Slice thickness: 4 mm | Abdomen | | | | |
| van Ufford et al. ³⁵ | 1.5 T | T1/T2 STIR: built-in body coil DWI: surface coil | T1 | Slice thickness: 6 mm | Coronal | Head, neck, chest, abdomen, pelvis |
| | | | T2 STIR | Slice thickness: 6 mm | | |
| | | | DWI | Slice thickness: 4 mm b-value: 0 and 1000 s/mm ² ADC not calculated | Axial | |
| Lin et al. ³⁶ | 1.5 T | Surface | DWI | Slice thickness: 5 mm b-value: 50, 400 and 800 s/mm ² ADC calculated | Axial | Head, neck, chest, abdomen, pelvis, proximal thigh |

SPAIR = spectral presaturation attenuated inversion recovery; STIR = short tau inversion recovery; DWI = diffusion-weighted imaging; SAT = fat saturation; ADC = apparent diffusion coefficient; MIP = maximum intensity projection; WB-MRI = whole-body diffusion-weighted magnetic resonance imaging.

WB-MRI and FDG-PET/CT in 105 cases (90.5%). In nearly all cases of differences in staging, WB-MRI yielded a higher grade than FDG-PET/CT. Overall, there was excellent agreement between the two methods ($\kappa = 0.871$; $P < 0.0001$).

Some characteristics inherent to WB-MRI may lead to false-positive results, such as its limited ability to distinguish malignant from benign causes of lymph node enlargement, particularly in inguinal and axillary nodes, and its extreme sensitivity for small lymph nodes, even on DWI sequences,¹⁴ as well as the T2 shine-through effect, which refers to an area of high signal on DWI mimicking restricted diffusion due to very prolonged spin-spin relaxation time.³⁴ In such cases, ADC quantitative analysis can be helpful for reducing the number of false positives in these cases. The causes of false-negative findings inherent to WB-MRI include diaphragmatic motion artifacts^{37,38} and artifacts in the hilar region due to respiratory and cardiac motion,³⁵ as well as falsely elevated ADC values in these areas.²⁵ Stéphane et al.¹⁶ also reported difficulties in analyzing hilar regions.

Several intrinsic factors may hinder interpretation of FDG-PET/CT results. Non-pathological variability in FDG uptake by healthy tissues, FDG uptake attributable to inflammation, altered biodistribution of FDG due to hyperglycemia or hyperinsulinemia and, particularly, the bone marrow activation commonly found in cancer patients after treatment may lead to false positives.³⁴

Addition of DWI to WB-MRI protocols provides improved lymph node viewing, compared with conventional sequences, thus increasing the accuracy of the method for detection of lesions,¹⁴ whereas ADC value analysis improves specificity. In the study by Lin et al.,³⁶ DWI using the lesion size criterion yielded sensitivity and specificity of 90% and 94% respectively, in comparison with FDG-PET/CT. Addition of visual ADC analysis reduced the sensitivity to 81% and increased the specificity to approximately 100%.

ADC quantification can also provide useful information on treatment response.¹⁸ In the study by Lin et al.,³⁶ the mean ADC ($\times 10^{-3}$ mm²/s) in regions with restricted diffusion was 0.75, versus 1.6 in regions with no restriction. In the study by Wu et al.,³⁴ the ADC correlated inversely with the maximum standardized uptake value (SUV_{max}) of FDG-PET/CT, thus suggesting that these parameters were comparable.

A different WB-MRI protocol was used in each of the six studies included, and this lack of standardization hindered comparison of WB-MRI with other, better established methods. Therefore, development of standardized protocols is critical to establishing the role of WB-MRI for staging and monitoring of lymphoma and other malignant conditions. Moreover, it may be of interest to compare different strength fields, such as 1.5 T versus 3.0 T. Despite this heterogeneity, five of the six studies

included showed excellent agreement between WB-MRI and FDG-PET/CT, and one of them showed good agreement.³⁵ Abdulqadhr et al.¹⁵ obtained T1-weighted and T2-weighted coronal images with SPAIR and axial DWI images, with 3D MIP reconstruction with a total scan time of roughly 50 minutes. Stéphane et al.¹⁶ used a similar protocol, except for the use of T2-weighted images with STIR, acquired for a total scan time of 40 minutes. Gu et al.¹⁴ used axial T2 and T2 with SPAIR sequences and axial DWI with 3D reconstruction for a total scan time of 44 to 52 minutes. Lin et al.³⁶ also obtained excellent results with axial DWI alone and a total scan time of 30 to 45 minutes. This study also included ADC calculation, thus improving the specificity of the method, which is essential for treatment response assessment. Wu et al.³⁴ also measured ADC and obtained results that corroborate its importance in patient follow-up, but the test protocol was complex and no information on total scan time was provided. van Ufford et al.³⁵ used a protocol consisting of coronal T1-weighted, T2-weighted and STIR images (total scan time, 25 to 30 minutes) and axial DWI (total scan time, 20 to 25 minutes). This was the only study in which WB-MRI showed good agreement with FDG-PET/CT, and this was due to lack of experience in WB-MRI interpretation by the examining radiologists. In our opinion, a WB-MRI protocol can be built only with DWI, which has shown excellent results for lymphoma staging in comparison with FDG-PET/CT. ADC analysis, visual or otherwise, should also be provided for, since assessment of the functional evolution of residual lesions plays an important role in treatment monitoring.³⁹

Diagnostic accuracy studies usually assess the accuracy of a test method under evaluation (index test) in relation to that of a gold-standard, well-established comparison method (reference standard), for detection of the presence or absence of a target condition. Conversely, the present review did not set out to assess the ability of WB-MRI or FDG-PET/CT to detect the presence or absence of lymphoma, but the ability of either method to yield a correct disease stage, in comparison with a reference standard. The clinical and radiological reference standard to which the results of WB-MRI and FDG-PET/CT were independently compared was based on a set of parameters assessed over time. Since this reference standard establishes the definitive baseline staging that will be used for patient management and treatment planning, it may be considered to be the true measurement. Therefore, we were able to calculate the sensitivity of the index methods as used in each of the studies included and compare them with the reference standard, i.e. to ascertain the ability of each method to stage the target condition correctly in relation to a true measurement.⁴⁰ We found that both WB-MRI and FDG-PET/CT exhibited high sensitivity in the studies by Gu et al.,¹⁴ Abdulqadhr et al.¹⁵ and Lin et al.,³⁶ ranging from 88 to 100% for

WB-MRI and 90 to 94% for FDG-PET/CT. The highest sensitivity (100% for both methods) was found in the study by Stéphane et al.¹⁶ In the study by Wu et al.,³⁴ because of a poorly representative patient spectrum and because both methods staged three out of the eight patients incorrectly, the overall sensitivity was 63%. In the study by van Ufford et al.,³⁵ the sensitivity of WB-MRI was 59%, and that of FDG-PET/CT, 73%.

One limitation of the present review derives from the use of a clinical and radiological reference standard. Stéphane et al.¹⁶ used FDG-PET/CT as the gold standard method, although they also explicitly used clinical and imaging follow-up data to set up the differences between WB-MRI and FDG-PET/CT. Gu et al.¹⁴ also referred FDG-PET/CT as the reference standard for assessment of lesions on an individual basis and established the Ann Arbor staging using data such as physical examination, integrated FDG-PET/CT images at baseline and follow-up, and bone marrow biopsy results. Both Abdulqadhr et al.¹⁵ and van Ufford et al.³⁵ separately staged the patients using WB-MRI and FDG-PET/CT. For the former, differences in staging between the two methods were resolved using biopsy results and clinical and CT follow-ups; for the latter, these differences were resolved using the contrast-enhanced full-dose component of the FDG-PET/CT examination and bone marrow biopsy. Lin et al.³⁶ staged patients by means of physical examination, contrast-enhanced CT, FDG-PET/CT and bone marrow biopsy. Neither van Ufford et al.³⁵ nor Lin et al.³⁶ made it clear whether follow-up examinations were also included in determining the final staging. Wu et al.³⁴ established lymphoma staging through detailed medical history, physical examination, standard laboratory tests, CT scans of the chest, abdomen and pelvis and bone marrow biopsy. All authors except Wu et al.³⁴ included WB-MRI and/or FDG-PET/CT as part of the reference standard and, because of this, incorporation bias may have occurred, which would probably increase the level of agreement between the two index tests and the reference standard, and hence overestimate the measurements of diagnostic accuracy.¹⁹

Proven presence or absence of viable tumor tissue in anatomical pathology specimens is the most accurate reference standard in the field of oncology. However, since lymphomas often present as a diffuse disease, surgical exploration of all potential sites of involvement for histological analysis is ethically and practically unfeasible and may not affect treatment planning; therefore clinical and radiological staging is widely accepted as the reference standard.

WB-MRI provides several advantages over FDG-PET/CT. It does not emit ionizing radiation, which is particularly useful in children and young adults^{41,42} and when patients must undergo repeated imaging for follow-ups, as in lymphoma cases. FDG-PET/CT exposes patients to substantial radiation doses and,

consequently, is associated with increased risk of later malignancies.⁴² Furthermore, thorough patient preparation is required before FDG-PET/CT, and because a cyclotron is required to produce FDG, it is not widely available.⁴³

CONCLUSION

WB-MRI is a highly sensitive method for initial lymphoma staging. It has excellent agreement with FDG-PET/CT and is a great alternative for managing lymphoma patients, without using ionizing radiation or an intravenous contrast agent. However, in order to define the role of WB-MRI in clinical practice, further studies are needed to assess the performance of WB-MRI in comparison with FDG-PET/CT, with regard to early and late response evaluation.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin.* 2012;62(1):10-29.
2. Armitage JO. Staging non-Hodgkin lymphoma. *CA Cancer J Clin.* 2005;55(6):368-76.
3. Connors JM. State-of-the-art therapeutics: Hodgkin's lymphoma. *J Clin Oncol.* 2005;23(26):6400-8.
4. Lister TA, Crowther D, Sutcliffe SB, et al. Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *J Clin Oncol.* 1989;7(11):1630-6.
5. Lapela M, Leskinen S, Minn HR, et al. Increased glucose metabolism in untreated non-Hodgkin's lymphoma: a study with positron emission tomography and fluorine-18-fluorodeoxyglucose. *Blood.* 1995;86(9):3522-7.
6. Schöder H, Noy A, Gönen M, et al. Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma. *J Clin Oncol.* 2005;23(21):4643-51.
7. Juweid ME, Cheson BD. Role of positron emission tomography in lymphoma. *J Clin Oncol.* 2005;23(21):4577-80.
8. Tsushima Y, Takano A, Taketomi-Takahashi A, Endo K. Body diffusion-weighted MR imaging using high b-value for malignant tumor screening: usefulness and necessity of referring to T2-weighted images and creating fusion images. *Acad Radiol.* 2007;14(6):643-50.
9. Jerusalem G, Beguin Y, Najjar F, et al. Positron emission tomography (PET) with 18F-fluorodeoxyglucose (18F-FDG) for the staging of low-grade non-Hodgkin's lymphoma (NHL). *Ann Oncol.* 2001;12(6):825-30.
10. Brenner D, Elliston C, Hall E, Berdon W. Estimated risks of radiation-induced fatal cancer from pediatric CT. *AJR Am J Roentgenol.* 2001;176(2):289-96.
11. Kleinerman RA. Cancer risks following diagnostic and therapeutic radiation exposure in children. *Pediatr Radiol.* 2006;36 Suppl 2:121-5.

12. Mathews JD, Forsythe AV, Brady Z, et al. Cancer risk in 680,000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. *BMJ*. 2013;346:f2360.
13. Walker RE, Eustace SJ. Whole-body magnetic resonance imaging: techniques, clinical indications, and future applications. *Semin Musculoskelet Radiol*. 2001;5(1):5-20.
14. Gu J, Chan T, Zhang J, et al. Whole-body diffusion-weighted imaging: the added value to whole-body MRI at initial diagnosis of lymphoma. *AJR Am J Roentgenol*. 2011;197(3):W384-91.
15. Abdulqadhr G, Molin D, Aström G, et al. Whole-body diffusion-weighted imaging compared with FDG-PET/CT in staging of lymphoma patients. *Acta Radiol*. 2011;52(2):173-80.
16. Stéphane V, Samuel B, Vincent D, et al. 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(11):2011-7.
17. de Bazelaire C, de Kerviler E. From multislice CT to whole-body biomarker imaging in lymphoma patients. *Eur Radiol*. 2011;21(3):555-8.
18. Dudeck O, Zeile M, Pink D, et al. Diffusion-weighted magnetic resonance imaging allows monitoring of anticancer treatment effects in patients with soft-tissue sarcomas. *J Magn Reson Imaging*. 2008;27(5):1109-13.
19. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-36.
20. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-74.
21. Lin C, Itti E, Luciani A, et al. Whole-body diffusion-weighted imaging with apparent diffusion coefficient mapping for treatment response assessment in patients with diffuse large B-cell lymphoma: pilot study. *Invest Radiol*. 2011;46(5):341-9.
22. Punwani S, Taylor SA, Bainbridge A, et al. Pediatric and adolescent lymphoma: comparison of whole-body STIR half-Fourier RARE MR imaging with an enhanced PET/CT reference for initial staging. *Radiology*. 2010;255(1):182-90.
23. Ribrag V, Vanel D, Lebouleux S, et al. Prospective study of bone marrow infiltration in aggressive lymphoma by three independent methods: whole-body MRI, PET/CT and bone marrow biopsy. *Eur J Radiol*. 2008;66(2):325-31.
24. Kwee TC, Fijnheer R, Ludwig I, et al. Whole-body magnetic resonance imaging, including diffusion-weighted imaging, for diagnosing bone marrow involvement in malignant lymphoma. *Br J Haematol*. 2010;149(4):628-30.
25. Kwee TC, van Ufford HM, Beek FJ, et al. Whole-body MRI, including diffusion-weighted imaging, for the initial staging of malignant lymphoma: comparison to computed tomography. *Invest Radiol*. 2009;44(10):683-90.
26. Brennan DD, Gleeson T, Coate LE, et al. A comparison of whole-body MRI and CT for the staging of lymphoma. *AJR Am J Roentgenol*. 2005;185(3):711-6.
27. Krohmer S, Sorge I, Krause A, et al. Whole-body MRI for primary evaluation of malignant disease in children. *Eur J Radiol*. 2010;74(1):256-61.
28. Stecco A, Romano G, Negru M, et al. Whole-body diffusion-weighted magnetic resonance imaging in the staging of oncological patients: comparison with positron emission tomography computed tomography (PET-CT) in a pilot study. *Radiol Med*. 2009;114(1):1-17.
29. Goo HW, Choi SH, Ghim T, Moon HN, Seo JJ. Whole-body MRI of paediatric malignant tumours: comparison with conventional oncological imaging methods. *Pediatr Radiol*. 2005;35(8):766-73.
30. Daldrop-Link HE, Franzius C, Link TM, et al. Whole-body MR imaging for detection of bone metastases in children and young adults: comparison with skeletal scintigraphy and FDG PET. *AJR Am J Roentgenol*. 2001;177(1):229-36.
31. Li S, Xue HD, Li J, et al. Application of whole body diffusion weighted MR imaging for diagnosis and staging of malignant lymphoma. *Chin Med Sci J*. 2008;23(3):138-44.
32. Kellenberger CJ, Miller SF, Khan M, et al. Initial experience with FSE STIR whole-body MR imaging for staging lymphoma in children. *Eur Radiol*. 2004;14(10):1829-41.
33. Iizuka-Mikami M, Nagai K, Yoshida K, et al. Detection of bone marrow and extramedullary involvement in patients with non-Hodgkin's lymphoma by whole-body MRI: comparison with bone and 67Ga scintigraphies. *Eur Radiol*. 2004;14(6):1074-81.
34. Wu X, Kellokumpu-Lehtinen PL, Pertovaara H, et al. Diffusion-weighted MRI in early chemotherapy response evaluation of patients with diffuse large B-cell lymphoma—a pilot study: comparison with 2-deoxy-2-fluoro- D-glucose-positron emission tomography/computed tomography. *NMR Biomed*. 2011;24(10):1181-90.
35. van Ufford HM, Kwee TC, Beek FJ, et al. Newly diagnosed lymphoma: initial results with whole-body T1-weighted, STIR, and diffusion-weighted MRI compared with 18F-FDG PET/CT. *AJR Am J Roentgenol*. 2011;196(3):662-9.
36. Lin C, Luciani A, Itti E, et al. 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(8):2027-38.
37. Kwee TC, Takahara T, Ochiai R, Nievelstein RA, Luijten PR. Diffusion-weighted whole-body imaging with background body signal suppression (DWIBS): features and potential applications in oncology. *Eur Radiol*. 2008;18(9):1937-52.
38. Kwee TC, Takahara T, Ochiai R, et al. Whole-body diffusion-weighted magnetic resonance imaging. *Eur J Radiol*. 2009;70(3):409-17.
39. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25(5):579-86.
40. Cooper KL, Harnan S, Meng Y, et al. Positron emission tomography (PET) for assessment of axillary lymph node status in early breast cancer: A systematic review and meta-analysis. *Eur J Surg Oncol*. 2011;37(3):187-98.

41. Brenner DJ, Elliston CD. Estimated radiation risks potentially associated with full-body CT screening. *Radiology*. 2004;232(3):735-8.
42. Huang B, Law MW, Khong PL. Whole-body PET/CT scanning: estimation of radiation dose and cancer risk. *Radiology*. 2009;251(1):166-74.
43. Kwee TC, Kwee RM, Nievelstein RA. Imaging in staging of malignant lymphoma: a systematic review. *Blood*. 2008;111(2):504-16.

This article was presented in the form of a dissertation by the author Rodrigo Regacini on October 18, 2012, to Universidade Federal de São Paulo-Escola Paulista de Medicina (Unifesp-EPM), São Paulo, Brazil

Sources of funding: None

Conflict of interest: None

Date of first submission: January 31, 2014

Last received: September 26, 2014

Accepted: October 28, 2014

Address for correspondence:

Rodrigo Regacini

Rua Napoleão de Barros, 800

Vila Clementino — São Paulo (SP) — Brasil

CEP 04024-002

Tel. (+55 11) 5908-7900

E-mail: regacini@gmail.com