ABSTRACT

Background: Correct staging and treatment initiation in malignant lymphoma depends on accurate lymph node characterization. However, nodal assessment based on conventional and diffusion-weighted (DWI) magnetic resonance imaging, remains challenging, particularly in smaller nodes.

Purpose: To evaluate first order apparent diffusion coefficient (ADC) texture parameters compared to mean ADC for lymph node characterization in Non-Hodgkin lymphoma (NHL) using whole-body DWI (WB-DWI).

Study type: Retrospective

Population: 28 patients with NHL

Field strength/sequence: 3 Tesla whole-body DWI using 2 b-values (0-1000 s/mm^2)

Assessment: Regions of interest were drawn on the 3 most hyperintense lymph nodes on b1000-images, irrespective of size, in all nodal body regions. Diagnostic performance of mean ADC (ADC_{mean}) was compared with first order ADC texture parameters: standard deviation (ADC_{stdev}), kurtosis (ADC_{kurt}) and skewness (ADC_{skew}). Additional subanalyses focused on accuracy of ADC_{mean} and ADC texture parameters in different lymph node volumes and nodal regions.

Statistical tests: Benign and malignant nodes were compared using Mann-Whitney-U-tests with FDG-PET/CT and bone marrow biopsy as reference standard. Receiver-Operating-Characteristic analyses were performed to determine cut-off values and calculate sensitivity, specificity, accuracy, positive and negative predictive value (PPV, NPV).

Results: ADC_{mean} ($p=0.008$), ADC_{skew} and ADC_{kurt} differed significantly between benign and malignant nodes ($p<0.001$), while ADC_{stdev} didn't ($p=0.21$). ADC_{skew} was the best discriminating parameter with 79% sensitivity, 86% specificity, 83% accuracy, 85% PPV and

81% NPV. In every volume category, ADC_{skew} yielded the highest accuracy (88% in 0-25th percentile volume, 75% in 25th-75th percentile, 93% in 75-100th percentile). On a per-region basis, ADC_{skew} accuracy varied 13.6% between nodal regions, while ADC_{mean}, ADC_{kurt} and ADC_{stdev} showed interregional variation of 17.4%, 20.3% and 14.9%, respectively.

Data conclusion: First order ADC texture analysis with WB-DWI improved lymph node characterization compared to ADC_{mean}. ADC_{skew} was the most accurate and robust discriminatory parameter over all lymph node volumes and nodal body regions.

Keywords: apparent diffusion coefficient, diffusion-weighted, MR imaging, lymphoma, staging

INTRODUCTION

Non-Hodgkin lymphoma (NHL) can arise virtually anywhere in the body, primarily in nodal tissue, and typically disseminates systemically via the lymphatics. Therefore, a whole-body imaging technique is required in order to cover all potentially affected regions. Currently, hybrid 18-Fluoro-deoxyglucose positron emission tomography computed tomography (18-FDG-PET/CT) is accepted as the imaging method of choice to stage aggressive NHL and FDG-avid indolent lymphoma[1]. However, FDG-PET/CT is cost-intensive, not always widely available, and is associated with a substantial amount of ionizing radiation.

Whole-body diffusion-weighted magnetic resonance imaging (WB-DWI) might serve as an alternative radiation-free imaging technique for lymphoma staging purposes[2–12]. Affected sites usually present as easily detectable masses, which are commonly hyperintense on the diffusion-weighted images acquired using a b-value of 1000 s/mm^2 (b1000-images) and hypointense on the calculated apparent coefficient diffusion (ADC) maps. However, smaller – even normal sized - nodes can also harbor malignant deposits and these might be difficult to identify as lymph nodes inherently exhibit high signal on b1000-images and low ADC values[13]. Furthermore, considerable ADC overlap exists between reactive and malignant lymph nodes, decreasing the technique's specificity[14].

First order texture analysis[15,16] is an emerging tool in MRI analysis, consisting of the determination of histogram derived parameters, which may provide insight on tumor heterogeneity. This approach has proven useful for differentiating between tumor types and histological grades[17–27] or assessing treatment response[28–30]. To date, texture analysis to discriminate benign and affected nodes has not been reported in malignant lymphoma.

Therefore, the aim of this study was to evaluate first order ADC texture parameters compared to the more commonly used mean ADC for lymph node characterization in NHL with WB-DWI with FDG-PET/CT as the reference of standard.

MATERIALS AND METHODS

Patients

This study was approved by the local ethics committee and all patients gave written informed consent prior to inclusion. Inclusion criteria were a (1) new diagnosis, of (2) histopathological proven (3) aggressive or indolent lymphoma. Exclusion criteria were defined as (1) general contraindications to MRI (e.g. pacemaker, claustrophobia) and (2) treatment initiation prior to the WB-DWI or between WB-DWI and FDG-PET/CT scans.

Twenty-eight patients (20 men and 8 women; age range 29-81 years, mean 60 years) with a new diagnosis of NHL (20 diffuse large B-cell lymphoma (DLBCL), 2 T-cell lymphoma (TCL) and 6 follicular lymphoma (FL)) were consecutively included. Patient characteristics are summarized in Table 1. Histopathological diagnosis and determination of subtypes were done according to the criteria of the current WHO classification of hematological and lymphoid malignancies[31] by an experienced pathologist.

All patients underwent WB-DWI in addition to routine clinical diagnostic procedures including physical examination, blood analysis, histological lymph node examination, bone marrow biopsy and FDG-PET/CT.

Chemotherapy is the standard treatment, although exact regimens vary between tumor types. Most common regimens in aggressive lymphoma are based on CHOP [cyclophosphamide, hydroxydaunorubicin (doxorubicin), vincristine, and prednisone/prednisolone] with or without Rituximab [R]; treatment of indolent lymphoma is generally based on CVP [cyclophosphamide, vincristine, and prednisone].

Imaging Techniques

WB-DWI

Whole-body imaging covering the brain down to the proximal 1/3rd of the upper legs was performed on a 3 Tesla (T) MRI system (Ingenia; Philips, Best, The Netherlands), with parallel radiofrequency transmission and phased-array head-neck, body and spine coils. Free-breathing short inversion time (TI) inversion recovery STIR WB-DWI was acquired in the transverse plane in 4 consecutive imaging stations (head/ neck, chest, upper abdomen and pelvis). Each imaging station consisted of 50 slices, 5 mm slice thickness, intersection gap of 0.1 mm, field of view (FoV) of 420 x 329 mm, voxel resolution of 4.6 x 4.7 x 5 mm, repetition time (TR)/TI of 8454/250 ms and echo time (TE) of 67 ms, with b-values of 0 (b0) and 1000 s/mm^2 (b1000). Total sequence time was 15 minutes and 21 seconds. For interpretation, coronal 5-mm thick whole-body multiplanar reconstruction (MPR) images were generated by the scanner software from the b0, and b1000 images as well as from the ADC-maps. For anatomical reference, whole-body STIR fat-suppressed T2-weighted turbospin echo (TSE) images were acquired in the coronal plane in 3 consecutive imaging stations. Each station consisted of 40 slices; 6 mm slice thickness, intersection gap of 0.6 mm, FoV of 263 mm x 228-452 mm, voxel resolution of 1.5 x 1.7 x 6 mm and TR/ TI of 2324-9294/200 ms and TE of 80 ms. Total sequence time was 6 minutes and 8 seconds. Finally, a sagittal T1-weighted TSE sequence was acquired over the spine in 2 consecutive stations (thoracic and lumbar spine) with the following parameters: 15 slices per station, 4 mm slice thickness, intersection gap of 0.4 mm, FoV 260 x 380 mm, voxel resolution of 1.3 x 1 x 4 mm and TR/TE of 378/7 ms. Total sequence time was 4 min and 24 s.

FDG-PET/CT

Patients fasted for 6 hours prior to the integrated FDG-PET/CT examination (Biograph 40 TruePoint with TrueV, Siemens Medical Solutions, Erlangen, Germany). FDG-PET-images were acquired 60 minutes after intravenous administration of FDG at an average dose of 302.6 MBq (range: 220-388 MBq) and in the same session a single-section, whole-body, spiral CT (40-slice Siemens Sensation, 85 mAs, 120 kV, slice thickness 5 mm, collimation 24 x 1.2 mm, table feed 23 mm/rotation) was performed after intravenous injection of 120 ml of a contrast agent containing 300 mg iodine/ml.

Image Analysis

Whole-body DWI

Two radiologists, (a board-certified radiologist with 10 years of experience (VV) and senior resident (KDP), blinded to all other study and clinical data analyzed the DW images in consensus using semi-automated contouring software (in-house developed MeVisLab (MeVis Medical Solutions AG, Bremen, Germany) program enabling tissue segmentation based on region growing). Lymph node regions were subdivided as follows: cervical (left/right), axillary (left/right), mediastinal and hilar, mesenteric, retroperitoneal (left/right), iliac (left/right) and inguinal (left/right). Lymph nodes were localized on the b1000 images. As small nodes can also have lymphomatous involvement, we chose not to use size criteria, but selected nodes based on their b1000 signal intensity (SI). Per region - both left and right when applicable - three lymph nodes appearing most hyperintense at b1000 images were selected for quantitative analysis. Each lymph node was consecutively annotated with the cursor and automatically contoured by the software, to form a 3-dimensional ROI per lymph node. In

case of mismatch between the ROI and nodal contour, the ROI contour was manually adapted to fit the nodal contour. ROIs were placed over the entire lymph node and in case of obvious solid and necrotic tissue on b1000 images ROIs were placed over the solid tissue portions. From these ROIs, pixel-wise ADC values were extracted for construction of an ADC histogram for each delineated lymph node, from which the mean (ADC_{mean}), standard deviation (ADC_{stdev}), kurtosis (ADC_{kurt}) and skewness (ADC_{skew}) were calculated.

Reference Standard Of Treatment Outcome

The histologic confirmation of the presence of lymphoma remains the reference standard. Routinely, a definite diagnosis of malignant lymphoma is made based on histopathological analysis of an excisional biopsy of an accessible node. However, as biopsy of every suspected lesion is practically and ethically not feasible, instead we used FDG-PET/CT, which was evaluated using the same nodal regional distribution as for WB-MRI/DWI, as an imperfect standard of reference to characterize the remainder of the lymph node regions. PET was performed as part of standard clinical care and reported by a nuclear medicine physician with 10 years of experience and a special interest in lymphoma, not blinded for clinical or other imaging data. A lymph node was considered PET-positive for lymphoma when FDG-uptake, marked higher than the background, was present in a location incompatible with normal anatomy or physiology. In case of discrepancies, follow-up imaging was used to characterize conspicuous lesions. Any lesion that decreased in volume during treatment was designated as malignant.

Statistical analysis

Statistical analysis was done with SPSS 13.0 for Windows (SPSS, Chicago, IL, US). A P-value < 0.05 was considered statistically significant.

Delineated lymph nodes on WB-DWI were correlated to FDG-PET/CT on a per-lymph node basis after all image interpretation was concluded. Descriptive values were calculated for every ADC parameter and ANOVA test with patient ID as random effect were conducted to check for significant differences between malignant and benign lymph nodes. ANOVA tests with Bonferroni correction were performed to investigate the possible confounding effect of factors such as disease stage, lymphoma subtype and lymph node volume on the ADC value. Next, Receiver-Operating-Characteristic (ROC) curves were constructed from which an optimal threshold equally weighting sensitivity and specificity for malignant lymph node detection was calculated. Subsequently, sensitivity, specificity, accuracy, negative and positive predictive value (NPV and PPV) were determined.

Additionally, the potential influence of nodal volume on parameter accuracy was investigated. For this purpose, lymph nodes were divided in 3 volume categories: the 25th percentile smallest nodes (≤p25), the 25th percentile largest nodes (≥p75) and all lymph nodes with a volume in between (p25-p75). Also, the potential influence of nodal body region on parameter accuracy was explored. As such, per-parameter Mann-Whitney U tests and ROC curve construction were repeated for every lymph node volume category and for the separate lymph node regions and associated accuracies were calculated.

RESULTS

First Order ADC Texture Parameters

All ADC histogram derived parameters demonstrated a significant difference between benign (n=140) and malignant nodes (n=101) (ADC_{mean}: F= 8.33, p=0.008; ADC_{skew}: F= 44.85, p<0.001; ADC_{kurt}: F= 15.15, p<0.001), except for ADC_{stdev} (F= 1.63, p= 0.21). ADC_{skew} yielded the highest AUC of 0.85 (Figure 1), attaining higher accuracy than the other first order ADC texture parameters, with a higher specificity of 86.4% and higher sensitivity of 79.2% (Table 2). A patient-by-patient analysis didn't reveal significant influence on any of the ADC values (ADC_{mean}: F= 0.55, p= 0.93; ADC_{skew}: F= 1.05 , p= 0.46; ADC_{kurt}: F= 1.07, p= 0.46). In addition, as demonstrated by post hoc correction, ADC values were not impacted by variations in stage (ADC_{mean}: F= 0.17, p= 0.84; ADC_{skew}: F= 0.81, p= 0.45; ADC_{kurt}: F= 1.91, p= 0.13), lymphoma subtype (ADC_{mean}: F= 0.91, p=0.34; ADC_{skew}: F= 1.81, p= 0.18; ADC_{kurt}: F= 0.083, p=0.77) or nodal volume (ADC_{mean}: F= 1.46, p=0.07; ADC_{skew}: F= 0.93, p= 0.64; ADC_{kurt}: F= 1.14, p= 0.31). Nodal region had a significant effect on ADC_{mean} (F= 7.70, p<0.001), but not on ADC_{kurt} (F= 2.15, p=0.049) or ADCskew (F= 1.42, p=0.21).

ADC_{mean} attained poor sensitivity of 58.4%, resulting in decreased NPV and accuracy. Inversely to ADC_{mean}, ADC_{kurt} demonstrated a higher sensitivity, but lower specificity with comparable accuracies of 65.8% and 67.2% for ADC_{kurt} and ADC_{mean}, respectively. Malignant nodes expressed lower ADC_{mean} (1.16 vs $1.43*10^{-3}$ mm^2/s), higher ADC_{kurt} (4.12 vs 3.22) and higher ADC_{skew} values (0.79 vs -0.002) than benign nodes (Figure 2).

First Order ADC Texture Parameter Analysis And Lymph Node Volume

Accuracy of ADC parameters in relation to lymph node volume is shown in Table 3. For all volume categories, ADC_{skew} was the best discriminating parameter. In the ≤p25 nodes (median volume [range] = 0.51 [0.10-0.65] cm^3), ADC_{skew} accuracy was 88.3%, which was 11%, 21% and 30% higher than the accuracies attained for ADC_{mean}, ADC_{kurt} and ADC_{stdev}, respectively. In the p25-p75 nodes (1.60 [0.66-3.69] cm^3), ADC_{skew} was the only significant parameter with an accuracy of 75.2%, 23-25% higher than the accuracy of the remainder of the first order texture parameters. In the ≥p75 nodes (62.97 [3.70-287.46] cm^3), all parameters except ADC_{mean} reached statistical significance with an ADC_{mean} accuracy of 68.3%, while the other parameters showed higher accuracies between 81.7-93.3%, the highest of which was again attained by ADC_{skew}.

First Order ADC Texture Parameter Analysis And Lymph Node Region

Using a per nodal region cut-off improved accuracy of ADC_{mean} and ADC texture parameters in the majority of lymph node regions, but also revealed interregional variations (Table 4). ADC_{skew} was the most robust parameter with an accuracy ranging between 78.1% and 91.4% (= total accuracy range of 13.3%) between lymph node regions. In relation to the overall accuracy, the use of a regional cut-off resulted in a 5.4% accuracy decrease for mesenteric and mediastinal nodes while the accuracy increased with 8% for inguinal nodes. ADC_{mean} variation of accuracy between regions was 17.4% with an accuracy increase of 4.2% in the iliac and 16.1% in the axillary nodes relative to the overall accuracy. ADC_{kurt} and ADC_{stdev} demonstrated an interregional accuracy variation of 20.3% and 14.9%, respectively. For both parameters, a regional cut-off resulted in higher accuracies in all lymph node regions of up to 23.3% for ADC_{kurt} in the mesenteric nodes and 15.4% for ADC_{stdev} in the inguinal nodes.

As indicated by the regional cut-off values, nodes in certain lymph node regions exhibited lower or higher ADC_{mean} and texture parameter values (Table 4, Figure 3). For ADC_{mean}, mediastinal and retroperitoneal nodes attained higher values in comparison with the other regions, especially compared to the axillary lymph nodes, which attained the lowest ADC_{mean} value. ADC_{skew} was lower in the mediastinal lymph node and higher in the iliac and inguinal region. Similarly, mediastinal and mesenteric nodes exhibited lower ADC_{kurt} values, while axillary and iliac nodes demonstrated a high ADC_{kurt}. Conversely, the highest ADC_{stdev} values were found for mediastinal and mesenteric nodes, and the lowest for inguinal nodes.

DISCUSSION

This study demonstrated that first order ADC texture analysis may improve lymph node characterization in malignant lymphoma compared to the use of ADC_{mean} - a commonly used, but unvalidated quantitative measure in lymphoma. ADC_{skew} proved to be the most accurate parameter to differentiate benign from lymphomatous nodes, and the most robust one in terms of lymph node size and lymph node region.

Quantitative DWI analysis of nodes has been mostly investigated in the context of differentiating metastatic, lymphomatous and benign cervical lymphadenopathy[32–38], and consisted of ADC_{mean} calculations with reported values of nodes involved with lymphoma ranging from $0.223 \pm 0.056 \times 10^{-3}$ mm^2/s to $0.97 \pm 0.27 \times 10^{-3}$ mm^2/s. The ADC_{mean} values demonstrated in lymphoma staging studies using a whole-body protocol[10,39,40] were within this range ($0.70 \pm 0.16 \times 10^{-3}$ mm^2/s to $0.87 \pm 0.17 \times 10^{-3}$ mm^2/s). In this study, a slightly higher ADC_{mean} of $1.09 \pm 0.28 \times 10^{-3}$ mm^2/s for malignant nodes was found, which might be explained by some methodological differences. In the aforementioned studies, ROIs were

manually drawn on calculated ADC maps, while in the current study a semi-automated delineation tool was used and ROIs were drawn on native DWI images. Also, the use of a scanner with different field strength – 3T study in the current study, and 1.5 T in previous literature -, may have resulted in absolute ADC value differences.

Using a ADC_{mean} cut-off value of 0.80 x 10^{-3} mm^2/s, Kwee et al.[41] yielded a sensitivity and specificity of 78% and 100% to detect lymphomatous nodes. Conversely, we found ADC_{mean} sensitivity of only 58.4% due to great overlap in ADC_{mean} values between benign and malignant nodes. Instead, first order texture analysis performed in this study demonstrated ADC_{skew} to be the parameter attaining highest sensitivity of 79.2% and specificity of 86.4%. This is in line with previous studies in other tumor types, showing the potential of this type of analysis for the discrimination of benign and malignant lesions. For instance, a significantly higher ADC_{skew} and lower ADC_{mean} were found in malignant endometrial[42] and breast lesions[43], while a study by Suo et al. concerning bladder lesions[44] showed ADC_{skew}, ADC_{kurt} and ADC_{mean} to differ significantly between benign bladder lesions and bladder carcinoma. Only the group of Wang et al.[45] has applied first order ADC texture analysis in malignant lymphoma to differentiate lymphoma from metastatic nodes in the head and neck region, but to our knowledge no other literature is available on first order texture analysis of lymphoma nodes using a whole body protocol.

The majority of studies using a quantitative analysis of WB-DWI to stage malignant lymphoma included enlarged nodes[4,6,10,40] only with a diameter of more than 10 mm. However, as normal sized nodes may contain malignancy as well, no size cut-off was implemented in our study. A subgroup analysis evaluating the diagnostic performance of ADC_{mean} and ADC texture parameters in different nodal volume categories demonstrated that ADC_{skew} was the most accurate parameter. In all volume categories ADC_{mean} attained rather

low accuracies ranging from 55% to 77%. In contrast to ADC_{skew}, the discriminatory capability of the remainder of the ADC texture parameters in middle-sized and small nodes was limited, although their attained accuracy in large nodes and nodal masses was good to excellent.

Besides accuracy variations in relation to nodal volume, variations were also noticed between lymph node regions. The use of regional cut-off values resulted in an increase of the diagnostic performance. Previous studies used an overall cut-off value based on either their own data [2,10,39,40] without the performance of a regional subgroup analysis or based on data from the literature[41], although these are mainly based on studies performed in cervical and/or pelvic nodes only. Yet, mediastinal and retroperitoneal nodes exhibited higher ADC_{mean} values than other nodal regions, which might be due to a variety of factors such as partial volume effect, motion artefacts (bowel, respiration) and distance to the coil. Although this has not been addressed before, it is of importance to avoid false negative and positive results when assessing individual nodal regions, and raises the question whether a regional cut-off should be used instead of an overall cut-off. The accuracy increase would probably be more noted for ADC_{mean}, a quite variable parameter in this study, than for a more robust parameter like ADC_{skew}.

This study had several limitations. First, quantitative parameters were used, the value of which is dependent of a multitude of factors (field strength, MRI scanner hard- and software, delineation, post-processing, etc.) and therefore prone for variations between observers and centers. In addition, the delineation of small nodes results in histograms with a small number of pixels, hence, small changes might induce large variations in measurements. We tried to overcome these limitations by applying a semi-quantitative delineation method, allowing for observer dependent node and ROI size selection, yet with automatic ROI delineation and

parameter calculations. This method also counteracts – at least partially- another important hurdle of quantitative analysis, being time-efficacy, which is currently limiting its implementation in daily practice. Furthermore, as lymphoma primarily present with large masses easily detected by visual WB-DWI analysis, a further reduction in analysis time might be achieved by performing quantitative analysis of equivocal nodes only. Finally, multiple confounding factors such as disease stage and lymphoma subtype as well as patient clustering may have affected our results. We tried to identify these factors by including corrections for multiple comparisons in the analysis. Yet, the small population is an important limiting factor, which might have hindered a meaningful statistical analysis. Nevertheless, our results were highly significant, warranting larger-scale studies to validate our findings. These larger datasets would not only allow for a more accurate evaluation of confounding effects, but should also explore the inter- and intra-observer variability. Of note, owing to the potential effect of data clustering in this limited dataset, the present optimal cut-offs and accuracies might differ from those found in larger studies.

In conclusion, first order ADC texture analysis using WB-DWI may improve lymph node characterization in malignant lymphoma. ADC_{mean}, although commonly used, demonstrated low accuracy, in contrast to $ADC_{skew,}$ which proved to be the best parameter to characterize lymph nodes regardless of lymph node volume or nodal body region.

REFERENCES

1. Barrington SF, Mikhaeel NG, Kostakoglu L, et al. 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*. 2014;32(27):3048-3058.
2. Mayerhoefer ME, Karanikas G, Kletter K, et al. 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(11):2984-2993.
3. Littooij AS, Kwee TC, Barber I, et al. Whole-body MRI for initial staging of paediatric lymphoma: prospective comparison to an FDG-PET/CT-based reference standard. *Eur Radiol*. 2014;24(5):1153-1165.
4. 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):2011-2017.
5. Lin C, Luciani A, Itti E, et al. Whole-body diffusion magnetic resonance imaging in the assessment of lymphoma. *Cancer Imaging*. 2012;12:403-408.
6. Gu J, Chan T, Zhang J, Leung AYH, Kwong Y-L, Khong P-L. 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.
7. 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):173-180.
8. Kwee TC, Takahara T, Vermoolen MA, Bierings MB, Mali WP, Nievelstein RAJ. Whole-body diffusion-weighted imaging for staging malignant lymphoma in children. *Pediatr Radiol*. 2010;40(10):1592-602-1.
9. Kwee TC, Vermoolen MA, Akkerman EA, et al. Whole-body MRI, including diffusion-weighted imaging, for staging lymphoma: comparison with CT in a prospective multicenter study. *J Magn Reson Imaging*. 2014;40(1):26-36.
10. Li S, Xue H-D, 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):138-144.
11. Balbo-Mussetto A, Cirillo S, Bruna R, et al. Whole-body MRI with diffusion-weighted imaging: a valuable alternative to contrast-enhanced CT for initial staging of aggressive lymphoma. 2016;71(3).
12. Albano D, Patti C, La Grutta L, et al. Comparison between whole-body MRI with diffusion-weighted imaging and PET/CT in staging newly diagnosed FDG-avid lymphomas. 2016;85(2).
13. Kwee TC, Takahara T, Luijten PR, Nievelstein R a J. ADC measurements of lymph nodes: inter- and intra-observer reproducibility study and an overview of the literature. 2010;75(2):215-220.
14. Chawla S, Kim S, Wang S, Poptani H. Diffusion-weighted imaging in head and neck cancers. *Future Oncol*. 2009;5(7):959-975.
15. Davnall F, Yip CSP, Ljungqvist G, et al. Assessment of tumor heterogeneity: an

emerging imaging tool for clinical practice? *Insights Imaging*. 2012;3(6):573-589.

16. Just N. Improving tumour heterogeneity MRI assessment with histograms. *Br J Cancer*. 2014;111(12):2205-2213.
17. Downey K, Riches SF, Morgan VA, et al. Relationship between imaging biomarkers of stage I cervical cancer and poor-prognosis histologic features: quantitative histogram analysis of diffusion-weighted MR images. *AJR Am J Roentgenol*. 2013;200(2):314-320.
18. Woo S, Cho JY, Kim SY, Kim SH. Histogram analysis of apparent diffusion coefficient map of diffusion-weighted MRI in endometrial cancer: a preliminary correlation study with histological grade. *Acta radiol*. 2014;55(10):1270-1277.
19. Ahn SJ, Choi SH, Kim Y-J, et al. Histogram analysis of apparent diffusion coefficient map of standard and high B-value diffusion MR imaging in head and neck squamous cell carcinoma: a correlation study with histological grade. *Acad Radiol*. 2012;19(10):1233-1240.
20. Pereira JAS, Rosado E, Bali M, Metens T, Chao S-L. Pancreatic neuroendocrine tumors: correlation between histogram analysis of apparent diffusion coefficient maps and tumor grade. *Abdom Imaging*. 2015;40(8):3122-3128.
21. Kim EJ, Kim SH, Park GE, et al. Histogram analysis of apparent diffusion coefficient at 3.0t: Correlation with prognostic factors and subtypes of invasive ductal carcinoma. *J Magn Reson Imaging*. 2015;42(6):1666-1678.
22. Takahashi M, Kozawa E, Tanisaka M, Hasegawa K, Yasuda M, Sakai F. Utility of histogram analysis of apparent diffusion coefficient maps obtained using 3.0T MRI for distinguishing uterine carcinosarcoma from endometrial carcinoma. *J Magn Reson Imaging*. 2016;43(6):1301-1307.
23. Lin Y, Li H, Chen Z, et al. Correlation of histogram analysis of apparent diffusion coefficient with uterine cervical pathologic finding. *AJR Am J Roentgenol*. 2015;204(5):1125-1131.
24. Brynolfsson P, Nilsson D, Henriksson R, et al. ADC texture--an imaging biomarker for high-grade glioma? *Med Phys*. 2014;41(10):101903.
25. Ryu YJ, Choi SH, Park SJ, Yun TJ, Kim J-H, Sohn C-H. Glioma: application of whole-tumor texture analysis of diffusion-weighted imaging for the evaluation of tumor heterogeneity. Hess CP, ed. *PLoS One*. 2014;9(9):e108335.
26. Rozenberg R, Thornhill RE, Flood TA, Hakim SW, Lim C, Schieda N. Whole-Tumor Quantitative Apparent Diffusion Coefficient Histogram and Texture Analysis to Predict Gleason Score Upgrading in Intermediate-Risk 3 + 4 = 7 Prostate Cancer. *AJR Am J Roentgenol*. 2016;206(4):775-782.
27. Hao Y, Pan C, Chen W, Li T, Zhu W, Qi J. Differentiation between malignant and benign thyroid nodules and stratification of papillary thyroid cancer with aggressive histological features: Whole-lesion diffusion-weighted imaging histogram analysis. *J Magn Reson Imaging*. 2016;44(6):1546-1555.
28. Liang H-Y, Huang Y-Q, Yang Z-X, Ying-Ding, Zeng M-S, Rao S-X. Potential of MR histogram analyses for prediction of response to chemotherapy in patients with colorectal hepatic metastases. *Eur Radiol*. 2016;26(7):2009-2018.
29. Kyriazi S, Collins DJ, Messiou C, et al. Metastatic ovarian and primary peritoneal

cancer: assessing chemotherapy response with diffusion-weighted MR imaging--value of histogram analysis of apparent diffusion coefficients. *Radiology*. 2011;261(1):182-192.

30. Breit A, Heuck A, Lukas P, Kneschaurek P, Mayr M. Tumor Response Monitoring and Treatment Planning: Advanced Radiation Therapy. 2013.
31. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20).
32. Fong D, Bhatia KSS, Yeung D, King AD. Diagnostic accuracy of diffusion-weighted MR imaging for nasopharyngeal carcinoma, head and neck lymphoma and squamous cell carcinoma at the primary site. *Oral Oncol*. 2010;46(8):603-606.
33. Sumi M, Van Cauteren M, Nakamura T. MR microimaging of benign and malignant nodes in the neck. *AJR Am J Roentgenol*. 2006;186(3):749-757.
34. Maeda M, Kato H, Sakuma H, Maier SE, Takeda K. Usefulness of the Apparent Diffusion Coefficient in Line Scan Diffusion-Weighted Imaging for Distinguishing between Squamous Cell Carcinomas and Malignant Lymphomas of the Head and Neck. *Am J Neuroradiol*. 2005;26(5).
35. Sumi M, Sakihama N, Sumi T, et al. Discrimination of metastatic cervical lymph nodes with diffusion-weighted MR imaging in patients with head and neck cancer. *AJNR Am J Neuroradiol*. 2003;24(8):1627-1634.
36. Holzapfel K, Duetsch S, Fauser C, Eiber M, Rummeny EJ, Gaa J. Value of diffusion-weighted MR imaging in the differentiation between benign and malignant cervical lymph nodes. *Eur J Radiol*. 2009;72(3):381-387.
37. King AD, Ahuja AT, Yeung DKW, et al. Malignant Cervical Lymphadenopathy: Diagnostic Accuracy of Diffusion-weighted MR Imaging. *Radiology*. 2007;245(3):806-813.
38. Abdel Razek AAK, Soliman NY, Elkhamary S, Alsharaway MK, Tawfik A. Role of diffusion-weighted MR imaging in cervical lymphadenopathy. *Eur Radiol*. 2006;16(7):1468-1477.
39. Wu X, Pertovaara H, Dastidar P, et al. ADC measurements in diffuse large B-cell lymphoma and follicular lymphoma: a DWI and cellularity study. *Eur J Radiol*. 2013;82(4):e158-64.
40. 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):2027-2038.
41. Kwee TC, Ludwig I, Uiterwaal CS, et al. ADC measurements in the evaluation of lymph nodes in patients with non-Hodgkin lymphoma: feasibility study. *Magn Reson Mater Physics, Biol Med*. 2011;24(1):1-8.
42. Kierans AS, Doshi AM, Dunst D, Popiolek D, Blank S V., Rosenkrantz AB. Retrospective Assessment of Histogram-Based Diffusion Metrics for Differentiating Benign and Malignant Endometrial Lesions. *J Comput Assist Tomogr*. 2016;40(5):723-729.
43. Suo S, Zhang K, Cao M, et al. Characterization of breast masses as benign or malignant at 3.0T MRI with whole-lesion histogram analysis of the apparent diffusion coefficient. *J Magn Reson Imaging*. 2016;43(4):894-902.

44. Suo S-T, Chen X-X, Fan Y, et al. Histogram analysis of apparent diffusion coefficient at 3.0 T in urinary bladder lesions: correlation with pathologic findings. *Acad Radiol*. 2014;21(8):1027-1034.

45. Wang Y-J, Xu X-Q, Hu H, et al. Histogram analysis of apparent diffusion coefficient maps for the differentiation between lymphoma and metastatic lymph nodes of squamous cell carcinoma in head and neck region. *Acta radiol*. September 2017:28418511773068.

TABLES

Table 1. Patient characteristics

	Aggressive NHL n=22	Indolent NHL n=6
Mean age (range) y	59 (17-83)	85 (48-75)
Sex		
Male	16 (72.7%)	4 (66.7%)
Female	6 (27.3%)	2 (33.3%)
Stage		
I	3 (13.6%)	1 (16.7%)
II	8 (36.4%)	0
III	4 (18.2%)	3 (16.7%)
IV	7 (31.8%)	4 (66.7%)
Subtype		
Grade1-2		5 (83.3%)
Grade 3a		1 (16.7%)
DLBCL	17 (77.3%)	
Burkitt	2 (9.1%)	
Primary mediastinal BCL	1 (4.5%)	
Peripheral TCL	2 (9.1%)	

DLBCL= diffuse B-cell lymphoma; BCL= B-cell lymphoma;
TCL= T-cell lymphoma

Table 2. Cut-off value and corresponding sensitivity, specificity, accuracy, negative predictive value and positive predictive value for mean ADC and ADC texture parameters

	ADC			
	Mean ($x10^{-3}$ mm^2/s)	Skewness	Kurtosis	Stdev ($x10^{-3}$ mm^2/s)
Cut-off	1.11	0.38	2.95	.22
Median [range] Ben (n=140)	1.34[.73-2.67]	.06[-1.61-1.45]	2.79[1.54-11.96]	.19[.03-.85]
Median [range] Mal (n=101)	1.04[.63-3.22]	.74[-1.14-2.05]	3.59[1.94-9.69]	.23[.03-.52]
Sensitivity (%)	58.4	79.2	76.2	55.4
Specificity (%)	73.6	86.4	57.9	61.4
Accuracy (%)	67.2	83.4	65.6	58.9
NPV (%)	71.0	85.2	77.1	65.7
PPV (%)	61.5	80.8	56.6	50.9

Ben= benign; Mal= malignant; Stdev= standard deviation;
NPV= negative predictive value; PPV= positive predictive value

Table 3. Accuracy of the ADC mean and texture parameters in volume-based subgroups

	<P25 (n=60) **mean [range] volume =** **0.51 [0.10-0.65] cm^3**			**P25-75 (n=121)** **mean [range] volume =** **1.60 cm^3[0.66-3.69] cm^3**			**>P75 (n=60)** **mean [range] volume =** **62.97 [3.70-287.46] cm^3**			
	Acc (%)	Median [range]		p-value	Acc (%)	Median [range]		p-value	Acc (%)	Median [range]		p-value
		Ben (n=54)	Mal (n=6)			Ben (n=80)	Mal (n=41)			Ben (n=6)	Mal (n=54)	
ADC_{mean}	77	1.37[.75-2.61]	.97[.74-2.57]	.082	55	1.28[.74-2.67]	1.20[.70-3.22]	.124	68	1.10[.73-1.66]	.94[.63-2.78]	.139
ADC_{skew}	88	.04[-1.61-1.29]	.90[0.26-1.18]	<.001	75	.11[-1.20-1.45]	0.59[-1.14-2.03]	<.001	93	.03[-.67-.28]	.91[-.17-2.05]	<.001
ADC_{kurt}	67	2.47[1.57-9.04]	3.09[2.62-3.33]	.139	52	3.03[1.72-11.96]	3.12[1.94-8.31]	.447	82	2.47[1.94-3.90]	4.23[2.40-9.69]	<.001
ADC_{stdev}	58	.16[.03-0.55]	.10[.03-.32]	.132	56	.20[.08-.85]	.22[.06-.52]	.358	88	.37[.25-.61]	.23[.11-.48]	.001

*ADC_{mean}= mean apparent diffusion coefficient (x 10^{-3} mm^2/s); stdev= standard deviation (x 10^{-3} mm^2/s), kurt= kurtosis, skew= skewness, Acc= accuracy, Ben= benign, Mal= malignant

*≤p25 equals the 25 percent smallest nodes, ≥p75 equals the 25 percent largest nodes and p25-75 equals all volumes in between.

Table 4. Per-region accuracies of all examined ADC parameters

		ADC			
		Mean ($x10^{-3}$ mm^2/s)	**Skew**	**Kurt**	**Stdev ($x10^{-3}$ mm^2/s)**
All	Cut-off	1.11	0.38	2.95	0.22
n=241	Acc	67.2%	83.4%	65.6%	58.9%
n=140	Med [range] Ben	1.34[.73-2.67]	.06[-1.61-1.45]	2.79[1.54-11.96]	.19[.03-.85]
n=101	Med [range] Mal	1.04[.63-3.22]	.74[-1.14-2.05]	3.59[1.94-9.69]	.23[.03-.52]
Ing	Cut-off	1.04	0.30	2.79	0.18
n=35	Acc	74.3%	91.4%	68.6%	74.3%
n=21	Med [range] Ben	1.16[.84-190]	.10[-.90-1.45]	2.69[1.57-4.81]	.13[.07-.23]
n=14	Med [range] Mal	.97[.63-1.67]	.99[.21-2.03]	3.37[2.58-8.31]	.20[.03-.48]
Il	Cut-off	1.10	0.63	4.00	0.16
n=35	Acc	71.4%	82.9%	71.4%	65.7%
n=21	Med [range] Ben	1.19[.76-2.16]	.21[-1.20-1.45]	3.24[1.98-8.85]	.15[.77-.32]
n=14	Med [range] Mal	1.02[.65-2.20]	1.16[-.52-2.05]	4.37[2.82-8.64]	.22[.06-.39]
Retro	Cut-off	1.42	0.46	2.90	0.20
n=38	Acc	81.6%	89.4%	71.1%	71.1%
n=26	Med [range] Ben	1.66[1.25-2.56]	-.05[-1.23-.75]	2.67[1.69-11.96]	.18[.11-.42]
n=12	Med [range] Mal	1.34[.79-3.22]	.84[-.39-1.89]	3.67[2.59-7.10]	.26[.18-.45]
Mes	Cut-off	1.30	0.35	2.54	0.24
n=9	Acc	77.8%	77.8%	88.9%	55.6%
n=4	Med [range] Ben	1.46[.92-1.73]	-.002[-.33-.35]	2.53[1.74-3.89]	.29[.19-.42]
n=5	Med [range] Mal	1.23[.73-1.66]	.56[.11-1.43]	3.31[2.65-5.34]	.26[.13-.37]
Med	Cut-off	1.60	0.28	2.45	0.23
n=32	Acc	78%	78.1%	78.1%	59.4%
n=12	Med [range] Ben	1.87[1.04-2.67]	.05[-.90-.54]	2.36[1.86-5.63]	.34[.15-.85]
n=20	Med [range] Mal	1.05[.75-2.78]	.56[-1.14-1.59]	3.69[1.99-6.35]	.24[.08-.39]
Ax	Cut-off	0.91	0.22	3.24	0.22
n=48	Acc	83.3%	87.5%	68.8%	62.5%
n=35	Med [range] Ben	1.15[.73-2.61]	-.01[-1.61-.86]	2.86[1.54-9.04]	.21[.05-.69]
n=13	Med [range] Mal	.85[.67-1.32]	1.10[-.55-1.95]	4.30[2.39-9.69]	.24[.10-.43]
Cerv	Cut-off	1.04	0.38	2.94	0.15
n=44	Acc	65.9%	81.8%	70.4%	61.4%
n=21	Med [range] Ben	1.21[.74-2.13]	.16[-1.33-1.29]	2.73[1.57-5.41]	.23[.03-.40]
n=23	Med [range] Mal	1.04[.72-1.70]	.66[.26-1.87]	3.40[1.94-7.77]	.19[.10-.52]

Stdev= standard deviation, Kurt= kurtosis, Skew= skewness; Acc= accuracy,
Ben= benign, Mal= malignant,
Ing= inguinal, Il= iliac, Retro= retroperitoneal, Mes= mesenteric, Med=mediastinal,
Ax= axillary, Cerv= cervical

FIGURE HEADINGS

Figure 1. ROC curve of mean ADC and ADC texture parameters standard deviation (stdev), kurtosis and skewness.

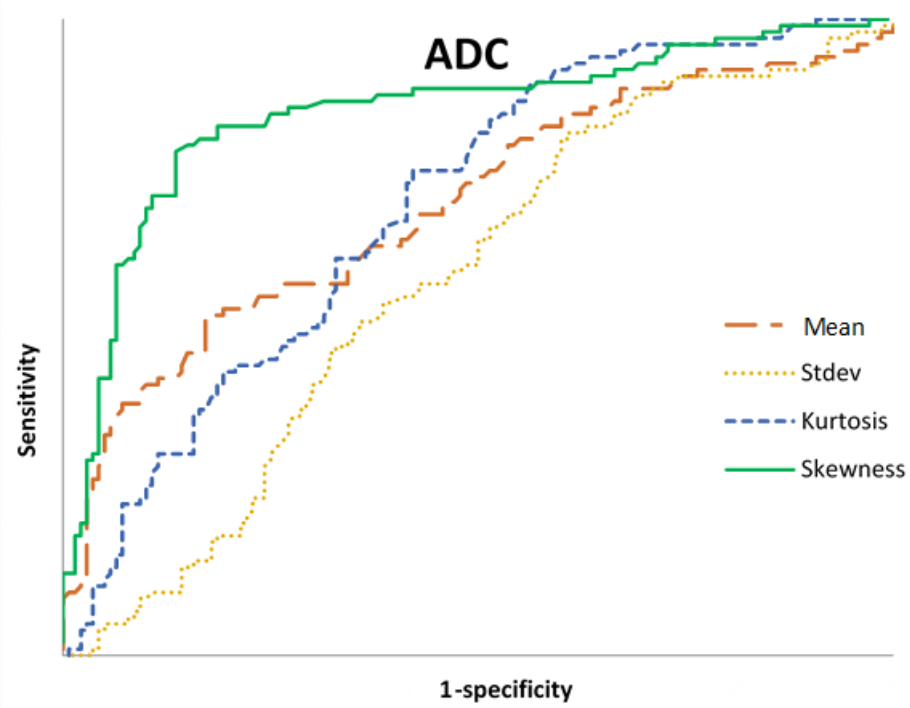

Figure 2. A) Illustration of a patient with stage II diffuse B-cell lymphoma, demonstrating a hyperintense malignant mass in the left cervical region (arrow) corresponding with an ADC histogram with an ADC_{mean} in the lower range, positive skew (high skewness) and a steep curve (high kurtosis). B) A different patient with a benign axillary lymph node, correlating with a higher ADC_{mean} value, a negative skew (low skewness), and a flatter shape (low kurtosis).

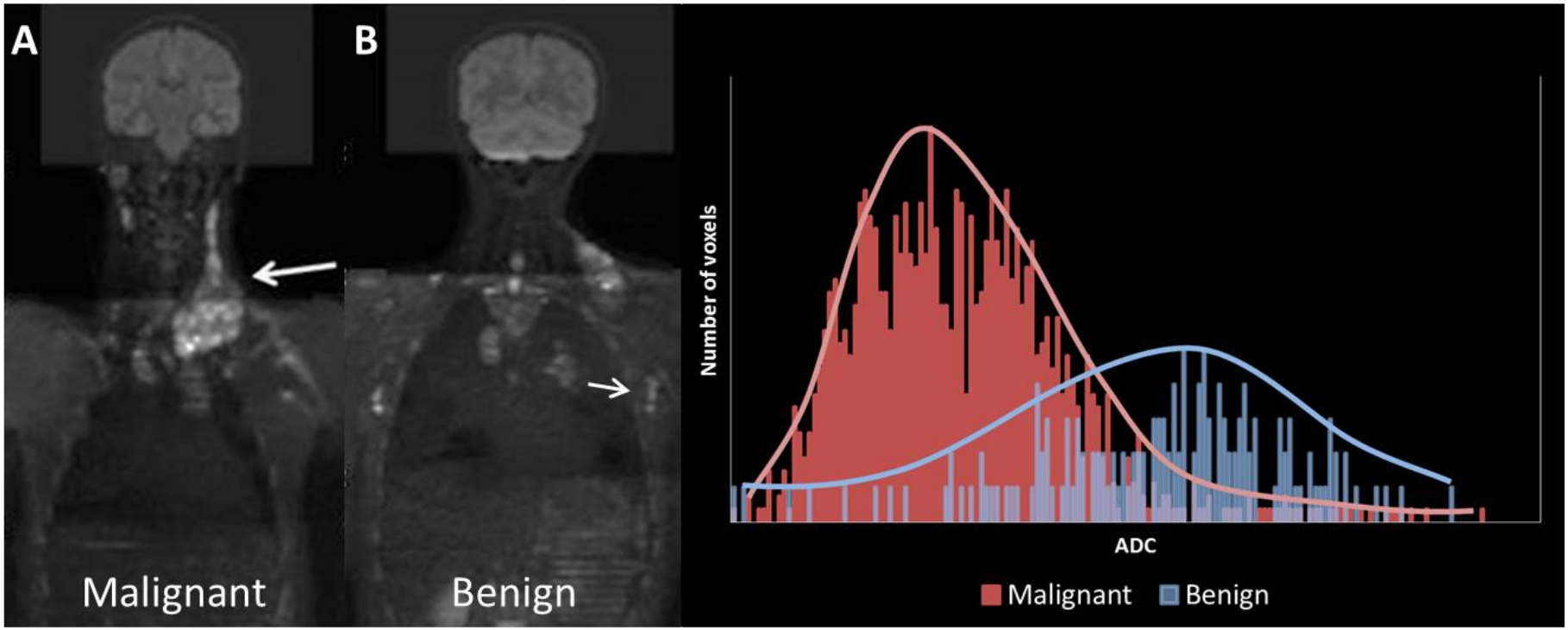

Figure 3. A) A 72-year old man with stage III Follicular lymphoma grade a.
B) Involved mediastinal nodes demonstrated a higher ADC mean, larger standard deviation, more negative skew, and lower kurtosis than axillary nodes and C) inguinal nodes.