Higher-order diffusion MRI characterization of mesorectal lymph nodes in rectal cancer

Andrada Ianuş[a,b*], Ines Santiago[c,d], Antonio Galzerano[c], Paula Montesinos[e], Nuno Loução[e], Javier Sanchez-Gonzalez[e], Daniel C. Alexander[b], Celso Matos[a,c], Noam Shemesh[a]

[a] *Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal,*
[b] *Centre for Medical Image Computing, University College London, London, UK*
[c] *Champalimaud Clinical Centre, Champalimaud Centre for the Unknown, Lisbon, Portugal,*
[d] *Nova Medical School, Campo Mártires da Pátria, 130, 1169-056 Lisbon, Portugal*
[e] *Philips Healthcare Iberia, Maria de Portugal 1. 28050, Madrid, Spain*

* *Corresponding author:*
Andrada Ianuş, Champalimaud Research, Champalimaud Centre for the Unknown, Lisbon, Portugal; Avenida Brasilia, 1400-038 Lisbon, Portugal;
e-mail: andrada.ianus@neuro.fchampalimaud.org;
phone number: +351 210 480 000 ext. #4467.

Word count: 5647

Running title: Higher-order diffusion MRI characterization of lymph nodes

Key words: lymph node imaging, rectal cancer, diffusion MRI, higher-order models, compartment models, IVIM, Kurtosis, diagnostic accuracy study

Abstract

Purpose: Mesorectal lymph node staging plays an important role in treatment decision-making. Here, we explore the benefit of higher-order diffusion MRI (dMRI) models accounting for non-Gaussian diffusion effects to classify mesorectal lymph nodes both 1) ex-vivo at ultrahigh field correlated with histology and 2) in-vivo in a clinical scanner upon patient staging.

Methods. The preclinical investigation included 54 mesorectal lymph nodes, which were scanned at 16.4T with an extensive dMRI acquisition. Eight diffusion models were compared in terms of goodness of fit, lymph node classification ability and histology correlation. In the clinical part of this study, 10 rectal cancer patients were scanned with dMRI at 1.5T and 72 lymph nodes were analysed with Apparent Diffusion Coefficient (ADC), Intravoxel Incoherent Motion (IVIM), Kurtosis and IVIM-Kurtosis.

Results. Compartment models including restricted and anisotropic diffusion improved the preclinical data fit as well as the lymph node classification compared to standard ADC. The comparison with histology revealed only moderate correlations, and the highest values were observed between diffusion anisotropy metrics and cell area fraction. In the clinical study, the diffusivity from IVIM-Kurtosis was the only metric showing significant differences between benign (0.80±0.30μm^2/ms) and malignant (1.02±0.41 μm^2/ms, p=0.03) nodes. IVIM-Kurtosis also yielded the largest area under the ROC curve (AUC=0.73) and significantly improved the node differentiation when added to the standard visual analysis by experts based on T2-weighted imaging.

Conclusions. Higher order diffusion MRI models perform better than standard ADC and may be of added value for mesorectal lymph node classification in rectal cancer patients.

Introduction

Characterizing the involvement of lymph nodes is a very important prognostic marker in rectal cancer patients as it determines the local recurrence and overall survival rates (1). To date, the gold standard for lymph node characterization is histology of the surgical specimen using hematoxylin and eosin (H&E) staining to highlight differences in cellular structure between normal lymphatic tissue and tumour (1). However, pre-operative characterization is beneficial because neoadjuvant chemoradiation has been shown to decrease the incidence of local relapse in high-risk patients (2), such as those with lymph node involvement. Thus, there is a clinical need for accurate pre-operative, preferably non-invasive, imaging techniques for lymph node characterization (3).

Magnetic Resonance Imaging (MRI) plays an important role in non-invasive imaging, including in cancer detection and staging (4). In the case of rectal cancer staging, T2-weighted MRI can be considered the contemporary gold standard for determining lymph node involvement. Unfortunately, T2 weighted MRI, whether based on size thresholds or morphologic characteristics such as shape, contour and heterogeneity has been shown to be quite a blunt tool with respect to lymph node characterization (5-7). Other contrasts, such as dynamic contrast enhancement (8), susceptibility contrast (9) or diffusion weighted imaging (10), have also been proposed for lymph nodes characterization with variable degrees of success.

Diffusion MRI (dMRI) is becoming an increasingly attractive modality, providing a non-invasive, indirect characterization of tissue microstructure (11, 12). The simplest dMRI metric, namely apparent diffusion coefficient (ADC), has been widely used for cancer imaging (13) as a biomarker reflecting increased cellularity in malignant tumours (14). For lymph node characterization, ADC has shown variable performance. For instance, some studies of cervical and axillary lymph nodes show that ADC is a promising technique for determining lymph node involvement (15, 16), while other studies, for example in cervical and perigastric lymph nodes, show a low accuracy of ADC for benign/malignant differentiation (17, 18). In rectal cancer, an accuracy of ~70% for lymph node involvement has been previously reported for ADC (19). As ADC assumes a single isotropic water pool which exhibits isotropic Gaussian diffusion, it lacks specificity to the underlying changes in tissue microstructure and cannot between variations in cell size, cellularity, orientation distribution, etc. This is also reflected in a variable correlation of ADC with histology data (14).

Higher-order dMRI models, which account for non-Gaussian diffusion, have been proposed to better characterize the diffusion properties in complex tissue microstructures. Diffusion Kurtosis (20) accounts for the signal departure from a mono-exponential decay at high b-values and has been successfully applied for cancer imaging, improving the differentiation of both primary tumours (21, 22) as well as lymph nodes (23-25), compared to ADC. To separate the effects of tissue perfusion and diffusion, the intra-voxel incoherent motion (IVIM) (26) captures the fast signal decay at low b-values due to flow effects, and, in many cases, provided better tumour characterization compared to ADC (27). For rectal cancer, IVIM also improved lymph node differentiation (28, 29), although the results are not consistent between the two studies. Combining IVIM and Kurtosis can provide insight into both perfusion and non-Gaussian diffusion effects, and has shown potential for differentiating various tumours (30, 31) as well as lymph nodes in the head and neck region (32), however, to our knowledge, it has not been employed so far to image lymph nodes involved in rectal cancer.

To further enhance the specificity of dMRI-derived parameters to the underlying microstructure, recently proposed higher-order dMRI techniques (33-40) employ biophysical compartment models of various (assumed) tissue features to fit the dMRI measurements, usually acquired at multiple b-values and diffusion times. Such approaches showed enhanced ability over ADC to explain the diffusion measurements and differentiate between benign and malignant tumours in various cancer types, such as xenograft colorectal tumours (33), prostate cancer (37, 41-44), breast cancer (40) and gliomas (45, 46). Moreover, the estimated dMRI parameters for restriction size and volume fraction correlated well with histology measurements (33, 41, 47). Compartment models have also been employed in a preliminary study to characterize diffusion in lymph node tissue ex-vivo (48), showing that models which include restriction and anisotropy provide the best fit to the data, although only 3 samples were included and a direct link to tissue microstructure through histological validation was not presented.

Here, we aimed to investigate whether a diffusion modelling approach could characterize lymph node structure and define lymph node involvement in rectal cancer. To this end, in the first part, we performed an ex-vivo study at ultrahigh field

MRI employing a rich dMRI protocol on lymph nodes extracted from the surgical specimens of rectal cancer patients. In this dataset, we compared various higher order diffusion models in terms of signal characterization, as well as node differentiation capability. We also correlated various diffusion metrics with quantitative histology features, which to our knowledge has not been studied so far. In the second part, we employed a clinically feasible dMRI acquisition and modelling approach at 1.5T to characterize higher order diffusion properties of mesorectal lymph nodes in rectal cancer patients upon staging, and investigated its benefit for node differentiation compared to standard T2-weighted qualitative assessment by expert radiologists.

Methods

Part 1: Ex-vivo imaging of lymph nodes at ultra-high field

The ex-vivo study aims to investigate the diffusion properties of benign and malignant lymph node tissue and their relationship with histological features in rectal cancer patients which were staged as N+, i.e. at least one lymph node has been classified as malignant following the pathological examination.

Institutional setting and lymph node harvesting

This study was approved by the institutional ethics committee and informed consent was obtained from all participating patients. Sixteen patients underwent surgery without neoadjuvant therapy and mesorectal specimens were excised and immersed in a 4% formaldehyde solution for 72h. Lymph nodes present in more than 1 cut slice (~ 5 mm thickness) were extracted, and processed as follows: one half was sent for pathologic staging performed by a gastrointestinal pathologist (8 years of experience), and the other half, which is normally discarded, was used for this study's ex-vivo scanning. Following the pathologic results, 5 additional patients with node-negative disease were excluded.

Prior to scanning, lymph node "halves" were washed with a 1% phosphate buffered saline solution for 24h and then mounted in a 10 mm NMR tube filled with Flourinert®, preferably grouped in pairs of similar size and originating from the same histopathologic block. In total, 75 nodes were scanned. Insufficient image quality due to susceptibility artifacts and/or image distortions retrospectively related to inadequate shimming, poor specimen preparation reflected by a highly inhomogeneous H&E staining in pathology as well as accidental sample mishandling resulted in the exclusion of 14 benign and 7 malignant nodes, as detailed in the Supporting Information Figure S1. Thus, the final analysis included 54 nodes (31 malignant) originating from 9 patients (mean age 63.2 years, 6 males).

Histopathologic analysis for comparison with MRI

To match histology to MRI, the scanned lymph nodes were embedded in paraffin and sliced parallel to cut surface every $50\mu m$ using a Leica RM2245 microtome (Leica Biosystems. Newcastle, United Kingdom). One slice per interval was stained with hematoxylin-eosin (H&E) and analyzed using a Zeiss Axio Lab A1, (Carl Zeiss Microscopy GmbH Germany) with a 40x amplification.

Diffusion MRI acquisition protocol at 16.4T

The ex-vivo MRI images were acquired on a 16.4T Bruker Aeon Ascend scanner interfaced with an AVANCE IIIHD console and operating a Micro5 probe with a gradient system capable of producing up to 3000 mT/m in all directions. A 10 mm birdcage coil was used for signal reception and RF transmission and samples were scanned at 37°C. The diffusion weighted images were acquired in stimulated echo acquisition mode (STEAM) using standard line-by-line readout with imaging parameters detailed in Table 1a. Single diffusion encoding (SDE) sequences with 6 gradient directions and duration $\delta = 1$ ms were acquired at four b-values up to 2000 s/mm^2 and multiple diffusion times Δ from 5 to 150ms, as detailed in Table 1a, leading to a total scan time of ~6h per node.

MR image analysis

Diffusion weighted images were analyzed in Matlab® (Mathworks, Natick, Massachussets, US). The acquired data was first normalized for each Δ to account for T_1 relaxation effects, due to different mixing times in the STEAM sequence. Then, eight diffusion models were fitted voxel-wise. The models included standard representations, such as ADC, diffusion tensor (DTI) and diffusion kurtosis (DKI), as well as several two-compartment models which account for anisotropic and/or restricted diffusion. In particular, the following models were considered (using the nomenclature in (49)): BallBall, BallSphere, ZeppelinBall, ZeppelinSphere, and BallFiniteCylinder, where “Ball” describes isotropic Gaussian diffusion (a single Ball compartment is equivalent to ADC), “Zeppelin” describes a cylindrically symmetric anisotropic Gaussian diffusion tensor, and “Sphere” describes isotropic, restricted diffusion. When fitting models with restriction, the same diffusivity for the restricted and Gaussian compartments was assumed (50). Details regarding the compartment models, specific parameter ranges and model assumptions are given in Supporting Information Table S1. For restricted diffusion, we used the signal expressions derived in (51) based on the Gaussian Phase Distribution approximation (52) for STEAM diffusion sequences (53-55).

Experiment 1 – Information-based comparison of DW-MRI models

The models were compared based on the Bayesian Information Criterion (BIC) which accounts for the goodness of fit and penalizes increasing number of model parameters:

$$BIC = ln(N)k - 2ln(\hat{L})$$

where N is the number of data points, k is the number of model parameters and $ln(\hat{L})$is the maximum value of the log likelihood function of the model, which accounts for Rician noise (51).

The BIC was computed voxel-wise for each model, then the models were compared based on the number of voxels where they occupy a certain rank.

Experiment 2 – Differentiation between benign and malignant nodes based on dMRI models

Next, we investigated the model parameter differences between benign and malignant lymph nodes for the standard DTI model as well as the best fitting model (ZeppelinSphere).

To study the ability of differentiating between benign and malignant nodes, we performed a receiver operating characteristic (ROC) analysis for the standard models (Ball, Kurtosis and Tensor) as well as for the best fitting ZeppelinSphere model. The parameters included in the analysis for each diffusion model were the following: Ball – MD, Kurtosis – MD and kurtosis, Tensor – MD and FA, ZeppelinSphere – f_{sphere}, D, $FA_{Zeppelin}$ and R_{sphere}. The ROC analysis was performed with two sets of histogram features for each model parameter: median values across each node and median plus inter-quartile range (IQR).

Experiment 3 – Correlation between quantitative histology and dMRI metrics

To investigate the link between parameters derived from dMRI data and histological features, we analyzed histology H&E slides and quantified the following parameters: cell count, average nuclear area, average cell area and cell area fraction. To ensure a relevant comparison between histology and MRI, the best matching H&E slice based on contour and visible landmarks was selected. Then, the histological features were quantified in a 420x560 μm^2 field using a cell detection tool (QuPath®, Belfast, UK).

A similarly sized ROI was placed in the dMRI images by a radiologist with 8 years of experience. Then, we quantified the correlation between the mean parameter values of the diffusion models and the histological features in the given ROIs. Specifically, we analysed the parameters of DTI and the best fitting ZeppelinSphere model. For the latter, we have excluded from the correlation analysis N = 97/1120 voxels where the estimated R_{sphere} reached the upper limit imposed during fitting, see Supporting Information Table S1.

Part 2: Clinical imaging at 1.5 T

The clinical study aims to investigate the benefit of including dMRI to differentiate between benign and malignant lymph nodes for rectal cancer patient staging.

Institutional setting

The clinical study was also carried out with the approval of the institution´s ethics committee. A total of 10 patients previously diagnosed with rectal cancer (mean age of 64.9 years, 5 males) were enrolled, after obtaining written informed consent. The imaging was performed on a 1.5T scanner (Ingenia, Philips Healthcare®, Best, The Netherlands) with the diffusion acquisition added to the clinical staging pelvic MRI. A spasmolytic agent was administered (Butylscopolamine 20 mg) and a pressure belt was utilized (Orfit Industries, Wijnegem, Belgium) to minimize bowel and respiratory movement artefacts, respectively.

Image acquisition and data analysis

The diffusion data was acquired using a Torso XL surface coil with multi-shot Spin-echo Echo-planar Imaging (SE-EPI), with the acquisition parameters detailed in Table 1b. The diffusion acquisition was split in three parts, with small b-values (50, 100, 200 s/mm^2) medium b-values (500, 1000 s/mm^2) and high b-values (1500, 2000, 2500 s/mm^2), with imaging parameters detailed Table 1b The sequences were split to minimize the echo time of the acquisition (65 ms), as previous visual optimization showed it plays an important role for image contrast. To achieve this echo time for the high b-values, the repetition time was increased to 9.5 s to satisfy duty cycle constraints. Thus, by separating the sequences, the repetition time for the small and medium b-values could be lower to allow an overall shorter acquisition time (TR of 3.5 and 3.7 for small and medium b-values respectively). Moreover, splitting the sequences allowed for multiple b0-values (in this case 3), which were beneficial for the delineation of the ROIs to minimize the effects of motion. The data was acquired in 3 orthogonal directions averaged on the scanner.

In total, 76 lymph nodes were delineated by a radiologist (8 years experience) and matched to pathology during macroscopy. The process consisted of sequential slicing of the specimen, followed by positioning and numbering of the cut slices to match the MR imaging acquisition and dyeing of the extracted lymph nodes based on their radial position. Due to motion artifacts, 4 nodes were excluded, leaving 72 nodes (14 malignant). Whole-node ROIs were defined on the high b-value images and copied to the low and medium b-value data sets. When necessary, the ROIs were slightly translated to account for motion during the acquisition. The average ROI signal for each b-value was normalized to the ROI average of the b0 images in the respective set. The normalized signal decay was fitted with four diffusion models, which can be used to characterize the diffusion properties of the tissue given the in-vivo clinical acquisition. Thus, ADC was fitted to the measurements with medium b-values, IVIM was fitted to the measurements with small and medium b-values, Kurtosis to the sets with medium and high b-values and IVIM-Kurtosis to the entire data set. In the clinical study, only models of isotropic diffusion without an explicit representation of restriction were fitted, as the acquired data did not have enough directional information for anisotropic models or diffusion times for models of restricted diffusion. On the other hand, the models included IVIM effects to account for perfusion.

We investigated the parameter differences between benign and malignant nodes and then employed a ROC analysis to compare the differentiation ability of the best performing model (IVIM + Kurtosis) with the standard T2-weighted classification performed by 2 experienced radiologists on a subset of 56 lymph nodes.

Statistical analysis

To investigate the parameter differences between benign and malignant nodes, we employed a mixed effect linear model with the diffusion metrics as response variable, and the malignancy status as predictor variable. The model accounts for the hierarchical structure of the data. In the voxel-wise analysis of the ex-vivo dataset, there is a three-level hierarchy, with the voxels grouped by lymph node and by patient. In the ROI analysis of the clinical data there is a two-level hierarchy, with the whole node ROI values grouped by patient.

To assess the differentiation ability of different models, we employ a ROC analysis. Specifically, we use the Matlab implemented following a multivariate linear regression. Differences between ROC curves were assessed using the DeLong test (56).

Results

Part 1: Ex-vivo imaging of lymph nodes at ultra-high field

Experiment 1 – Information-based comparison of dMRI models

The first experiment compared the goodness of fit of different diffusion models according to their BIC values. Figure 1a) illustrates the model ranking for benign (left) and malignant nodes (right). The models are listed in decreasing order of their performance from the top to the bottom of the plots. The results show that models including anisotropy and restriction consistently rank higher both for benign and malignant nodes. Figure 1b) presents maps of the best fitting model for a benign and a malignant node. A spatial variability of the best fitting model pattern can be observed, which reflects the underlying structure, with the ZeppelinSphere model providing the best fit in the nodal parenchyma, while the tensor model provides the best fit in the nodal capsule. This pattern is consistent for other nodes where the capsule is visible (data not shown).

Figure 2 shows example datasets for two voxels from a benign (left) and a malignant (right) lymph node which were fitted with the a) Tensor model and b)

ZeppelinSphere model. In both node types the ZeppelinSphere model fit (solid lines) better captured the trends in the experimental data (symbols) compared with the Tensor model, as expected from the results presented in Figure 1. The data also shows a higher spread of the measured signal in the diffusion time dimension for the malignant voxel compared to the benign voxel. The time-dependence pattern is variable across nodes; nevertheless, on average, malignant nodes exhibit stronger diffusion time dependence compared to benign nodes.

Experiment 2 – Differentiation between benign and malignant nodes based on dMRI models

The next analysis targeted the ability of different model parameters to differentiate between benign and malignant nodes. For the standard Diffusion Tensor model, when analyzing the voxelwise parameters from all nodes, both MD and FA show significant differences between benign and malignant nodes, with lower MD and higher FA in malignant nodes ($p<0.01$ in both cases), as illustrated in Supporting Information Figure S2. The results for the best fitting ZeppelinSphere model are presented in Figure 3. Parameter maps of diffusivity (D_i), fractional anisotropy of the zeppelin compartment (FA), fraction of spherical compartment (f_{sphere}) and radius of the spherical compartment (R_{sphere}) are shown in panel a) and the parameter differences between benign and malignant nodes in panel b). D_i shows a significant decrease between benign and malignant nodes ($p<0.01$), while FA and f_{sphere} show a significant increase ($p<0.01$). The sphere radius does not evidence any statistically significant difference between the groups.

Figure 4 presents the ROC curves of four diffusion models (Ball, Kurtosis, Tensor and ZeppelinSphere), based only on median parameter values (a) or median and IQR values (b). When analyzing only the median values, the areas under curve (AUC) for all models considered were between 0.7 and 0.8, with the ZeppelinSphere showing a higher AUC (0.79). However, this difference was not significant based on the DeLong test. When including the IQR of the parameters, a significant increase ($p<0.01$) in the classification ability of the Tensor and ZeppelinSphere models was observed (AUC of 0.93 and 0.95, respectively). We have further tested whether the Tensor model fitted to a single shell acquisition ($b = 1000$ s/mm^2, $\Delta = 5$ ms) could also provide a good differentiation between benign and malignant nodes. The ROC analysis showed similar values of the AUC (0.78 and 0.92 for median and median +

IQR, respectively) which, although slightly smaller, are not significantly different from the ZeppelinSphere results. These results, together with the fact that including the IQR for the ADC and BallSphere models does not result in a significant increase in the AUC (data not shown), suggest that both anisotropy and its heterogeneity in the malignant lymph nodes plays an important role for the classification.

Experiment 3 – Correlation between quantitative histology and dMRI metrics

Figure 5 summarizes the results of the quantitative histology analysis, based on corresponding ROIs drawn on microscopy and MRI images, as presented in Figure 5a) and b). The results show statistically significant differences ($p<<0.01$) between the benign and malignant lymph nodes for all the histological features included in the analysis (cellularity, average nuclear area, average cell area, cell area fraction).

Figure 6 shows the significant correlation between dMRI derived metrics and the histology features described above. The correlations for the Tensor parameters are depicted in Figure 6a). MD showed a weak positive correlation with cell count ($r=0.34$, $P=0.014$) and weak negative correlations with nuclear area ($r=-0.38$, $P=0.006$), cell area ($r=-0.39$, $P=3.4E-3$) and cell area fraction ($r=-0.36$, $P=0.011$). FA exhibited a moderate positive correlation with nuclear area ($r=0.57$, $P=1.5E-5$) and cell area fraction ($r=0.59$, $P=4.4E-6$). The correlations for the ZeppelinSphere model parameters are depicted in Figure 6b). Zeppelin FA showed a moderate positive correlation with nuclear area ($r=0.59$, $P=7.2E-6$) and with cell area fraction ($r=0.66$, $P=1.7E-7$). R_{sphere} exhibited a moderate positive correlation with cell count ($r=0.46$, $P=7.7E-4$), and a moderate negative correlation with cell area ($r=-0.44$, $P=1.6E-3$) and the corresponding cell radius ($r=-0.42$, $P=2.4E-3$). For D_i the correlation coefficients were smaller than 0.3. Another feature of this dataset is a large variability of the estimated parameters in the benign lymph nodes, especially for the diffusivities and sphere radius, which has a direct impact on the observed correlations.

Part 2: Clinical imaging at 1.5T

This section presents the analysis of the clinical dMRI data acquired at 1.5T. In total 72 lymph nodes, originating from 10 patients with rectal cancer and without neoadjuvant therapy, were analyzed.

Figure 7 a) and b) illustrate examples of diffusion weighted images and ROI selection for a benign and a malignant node, respectively, and Figure 7c) plots the average signal decay separated by the malignancy types. The decay curves in Figure 7c) clearly revealed a faster decay for malignant nodes and a slower decay for benign nodes, which are also reflected in the model parameters presented in Figure 8 and Table 2.

Figure 8a) shows box-plots of the parameter values for ADC, IVIM, Kurtosis and IVIM-Kurtosis estimated in benign and malignant nodes. ADC values were higher in malignant nodes compared with benign nodes; however, the difference was not statistically significant (P=0.26). The only parameter which reached a statistically significant threshold was the diffusivity (D) estimated from the IVIM-Kurtosis model (P=0.03), which has higher values in the malignant nodes. The diffusivity values from the IVIM and Kurtosis models were also higher in malignant nodes, however the differences were not statistically significant (P=0.11 and 0.22 for IVIM and Kurtosis, respectively). The pseudo-diffusivities from the IVIM and IVIM-Kurtosis models are lower in malignant nodes compared to benign nodes, with p values of 0.07 and 0.24, respectively. The kurtosis values in malignant nodes were also lower than in benign nodes, however the difference was not statistically significant (P=0.12 for the Kurtosis model and P=0.32 for the IVIM-Kurtosis model).

Figure 8b) presents the ROC analysis which compared the four diffusion models in terms of differentiating between benign and malignant lymph nodes. ADC is the worst performing model, with an AUC value of 0.60, while IVIM-Kurtosis is the best performing one with an AUC of 0.73, however, the difference does not reach statistical significance according to the DeLong test (P=0.19). Figure 8c) shows that IVIM-Kurtosis also provides a better differentiation between benign and malignant nodes (AUC=0.80) compared to the standard clinical T2 classification based on the ESGAR 2016 criteria (57) (AUC=0.74 for Reader 1 and AUC=0.59 for Reader 2), which takes into consideration lymph node size, shape, contour and heterogeneity. Moreover, combining IVIM-Kurtosis and T2 results significantly improves the classification compared to T2 weighted images alone (P=0.08 for Reader 1 and P=0.007 for Reader 2).

4. Discussion and conclusions

Characterising lymph nodes in-vivo and noninvasively is of high relevance given the importance of lymph node staging in rectal cancer patient treatment planning (2). Since lymph node microstructural changes play an important role in malignancy (1) and given that diffusion MRI can portray at least some microsctructural properties, we hypothesized that dMRI could be used to distinguish malignant from benign nodes.

In the first part of this study, lymph nodes underwent extensive dMRI experiments at ultra-high field MRI, leading to a rich and robust dataset, with excellent resolution, to which eight different compartment models were fitted. An important finding of the pre-clinical study is that models including restriction and anisotropy best explain the measured data for both benign and malignant nodes (Figure 3), a result in agreement with preliminary data from (48). Spatial variations of the diffusion patterns are also observed. Thus, the lymph node capsule exhibited higher anisotropy and lower restriction compared with the lymph node parenchyma, with the diffusion tensor being the best fitting model. The tensor MD and FA are higher in the capsule compared to the parenchyma, as illustrated in Supporting Information Figure S4. This probably reflects the inherent capsule composition, which is relatively paucicellular (dominant cell being the fusiform fibroblast) and extracellular matrix-rich, with long extracellular matrix molecules, predominantly of collagen and elastin. Besides providing a better fit to the data, the higher-order models also improved the differentiation between benign and malignant nodes (Figure 4), better capturing the microstructural differences. While benign nodes are largely comprised of leucocytes, which are predominantly round, small and practically devoid of cytoplasm, malignant nodes predominantly contain large and goblet shaped malignant cells and exhibit a wide range of malignancy patterns including solid infiltration, necrosis and desmoplasia, with or without an associated inflammatory infiltrate. This wide range of cellular structure could give rise to high diffusion heterogeneity poorly captured by ADC.

Previous studies characterizing the diffusion properties of ex-vivo lymph nodes have mainly employed ADC (48, 58, 59), however, the results have shown variable outcome, with a very large spread in ADC values (48), as well as poor correlation between in-vivo and ex-vivo lymph nodes (58). Our study is consistent with these findings: the results evidenced a very large spread of ADC even for benign nodes (Figure 6). Potential sources for this high variability could include variable microstructure, fixation effects, or interactions with internal, susceptibility-induced fields. Regarding the latter, a large proportion of the lymph nodes analyzed in this study have also been imaged using a multi-gradient echo (MGE) acquisition (9), whose output (T2*) reflects the field inhomogenity. Similar to ADC, the T2* values from (9) also exhibited a large spread in the same benign nodes, however, no significant correlations have been found between ADC and T2* values, as illustrated in Supporting Information Figure S3. Regarding the effect of fixation duration, previous studies have shown variable results depending on the tissue type (59, 60),(61). In our study, the fixation interval ranged from 2 weeks to 35 weeks (median 15 weeks), however, the correlation between MD values (or other metrics) and fixation period ($r=0.33$, $P=0.09$) was not significant. Thus, the preparation cannot account for the full variability observed in the benign nodes.

Another interesting finding of the preclinical aspect of this study is that even the best-fitting model – despite its success in characterizing the diffusion signal decay and differentiating between malignant/benign nodes – correlates only moderately with the histological features. In malignant nodes, the restriction size r_{sphere} has similar values to cellular sizes in histology, with an average underestimation of 21±32%; however, for benign lymph nodes, the restriction size has an average overestimation of 203±110%, with a large spread of values, suggesting that the model is not necessarily adequate in benign lymph nodes. The highest correlation coefficient was observed between the FA values and the intracellular area fraction from histology, suggesting that the packing is somehow captured by the FA while the restriction effects fitted by the models average out other histological properties which are not directly visible in H&E staining and therefore correlate worse with the specific features they aim to represent, such as cell size.

Taken together, the above-mentioned features highlight the simplistic nature of tissue modelling in terms of *specificity* and calls for caution when interpreting the model parameters (e.g., the model's spherical radii assigned to cellular spaces). Indeed, several assumptions are made. For example, the models assume non-exchanging water pools. However, as the diffusion times were varied up to 150ms, exchange between compartments cannot be completely ruled out (62, 63), which potentially leads to an underestimation of the restriction fraction (50). Another assumption is having the same diffusivity inside and outside the sphere, with time-dependence originating only from restricted diffusion, which might once again lead to bias (46, 64). Nevertheless, this assumption has been successfully applied in the past to model various tumour types (33, 37, 40), with fitted parameters showing a good correlation with histology features (33, 41).

It is important to stress that, even if the estimated model parameters do not naturally correspond to features visible to histology, the modelling approach successfully captured the higher order effects in the signal decay and recovered parameters which were *useful* for lymph node classification, whereas the ADC approach was much less sensitive. In the future, more data could be acquired, thereby enabling increasingly complex models to be fit to the data, especially if single-shot acquisitions would be performed, leading to much more rapid acquisitions. Additionally, the sample preparation for ex-vivo scanning could be improved in future experiments. In this study some excluded nodes presented with poor quality images due to susceptibility artefacts which we related to the accumulation of blood degradation products at their cut surface as well as image distortions which we retrospectively related to imperfect shimming. We believe imaging intact lymph nodes and limiting the amount of surrounding fat by more attentive dissection might minimize susceptibility and distortion in future experiments, respectively.

The second part of this study aimed at investigating some of the diffusion properties described above for clinical characterization of mesorectal lymph nodes in patients with rectal cancer at 1.5T. Given the clinical constraints, modelling anisotropy and restriction was not feasible, and perfusion effects had to be considered. Nevertheless, higher-order effects, shown to be important in the pre-clinical study, were accounted for in Kurtosis models. The results show that ADC alone does not faithfully distinguish between malignant and benign nodes, whereas an increase in the area under the ROC curve was noted for the IVIM-Kurtosis model compared with the standard ADC (P=0.19). Moreover, combining IVIM-Kurtosis and T2 weighted images provided a significant increase over the standard clinical classification based on T2 weighted images alone (P=0.08 for Reader 1 and P<0.01 for Reader 2). These results show that it is important to disentangle the various signal contributions which are reflected in the ADC values in order to better characterize and distinguish between node types.

The diffusion parameters obtained in this analysis for the ADC, IVIM and Kurtosis models, i.e. higher ADC and D and lower D* and K in malignant nodes, are consistent with two recent studies employing IVIM and Kurtosis, respectively, to image lymph nodes in patients with rectal cancer (25, 28), however in contradiction with the IVIM results from (29), where the diffusivity in the malignant nodes is lower. The wide range of ADC and IVIM parameters are likely due to the heterogeneity of the lymph nodes as well as possible influences from the acquisition protocol. Moreover, in these studies there is no mention of any measures taken to minimize susceptibility and movement artifacts, which are common in pelvic MRI and whose avoidance was very important in our experiment.

Indeed, this study utilized state-of-the-art techniques to minimize motion artifacts, specifically, patients received a spasmolytic agent to minimize bowel movement and a pressure belt was used to minimize artifacts due to respiratory motion. Nevertheless, some motion artifacts and partial volume could still be detected, especially when considering such small structures as lymph nodes. To further reduce the impact of motion on the data analysis, the dMRI parameter quantification was performed on ROIs which were slightly translated for each part of the acquisition. Motion artifacts then lead to the exclusion of only 4 lymph nodes from the analysis, which we consider a good outcome for such MRI-challenging structures.

One limitation of this clinical study is the relatively small number of patients (N=10) and the fact that half of malignant lymph nodes (7/14) originate from only one patient. Nevertheless, this effect was accounted for in the statistical model. Another limitation is related to the diffusion times and b-values clinically achievable. Although, the pre-clinical results revealed that models with restriction and anisotropy best explain the diffusion properties of the tissue, such models were too complex to be fitted to the clinical data, which was constrained to measurements in three orthogonal directions averaged directly on the scanner and similar diffusion times for the different b-values. With the current imaging resolution, it was not feasible to assess the heterogeneity of diffusion parameters in the lymph nodes, with some nodes consisting only of a few voxels. Moreover, partial volume and motion effects, as well as different contrast between b0 and diffusion weighted images makes image registration a very difficult problem, thus the analysis was performed at ROI level which is more robust, but does not inform on parameter heterogeneity. Recent technical improvements in clinical scanners (slew rate, gradient strength, 3T field strength) would enable future studies to acquire dMRI measurements with multiple directions and diffusion times with a reasonable scan duration, which is required to characterize restriction and anisotropy in the tissue, as well as higher resolution images which could better inform about tissue heterogeneity. Nevertheless, these diffusion properties and their heterogeneity are to some extent reflected in the higher-order kurtosis models. Although the ex-vivo and clinical studies cannot be directly compared as they include different patients, drastically different field strengths and image resolution as well as different diffusion acquisitions and models, in both cases the results show that models which include higher order diffusion effects provide the best differentiation between benign and malignant nodes. Moreover, these approaches show a similar performance for the clinical data and ex-vivo data, when considering parameters averaged over the entire node.

To conclude, this work shows the potential of higher order dMRI models to characterize and differentiate benign from malignant lymph nodes in rectal cancer patients, both ex-vivo at ultrahigh field and in-vivo at 1.5T, paving the way for future oncologic studies on lymph node staging.

Funding: This study was funded by the Champalimaud Centre for the Unknown. Dr. Andrada Ianus´ and Prof. Daniel C. Alexander's work was also supported by EPSRC grants EP/M020533/1 and EP/N018702/1 and the NIHR UCLH Biomedical Research Centre. The contributions of Paula Montesinos, Nuno Loução and Javier Gonzalez-Sanchez were funded by Philips Healthcare Iberia.

Acknowledgements. The authors would like to thank the following for their contributions:

Lara Castanheira, Alexandra Ferreira and Ana Santos from the Pathology Department of the Champalimaud Centre for the Unknown;

Nuno Figueiredo and José Filipe Cunha from the Digestive Unit of the Champalimaud Centre for the Unknown;

Carlos Leichsenring, Vasco Geraldes and Cisaltina Sobrinho, from the Department of Surgery of the Hospital Fernando Fonseca, EPE;

Maria Lisitskaya, a fellow radiologist at the Computational Clinical Imaging Group, Champalimaud Centre for the Unknown;

Nickolas Papanikolaou from the Computational Clinical Imaging Group, Champalimaud Centre for the Unknown;

Eleftheria Panagiotaki from the Centre for Medical Image Computing, University College London;

All the staff from the Radiology Department of the Champalimaud Centre for the Unknown;

The researchers from the Neuroplasticity and Neural Activity MRI Lab, Champalimaud Centre for the Unknown.

References

1. Yao Y-F, Wang L, Liu Y-Q, et al. Lymph node distribution and pattern of metastases in the mesorectum following total mesorectal excision using the modified fat clearing technique. J Clin Pathol. . 2011;64(12):1073-7.
2. Sauer R, Becker H, Hohenberger W, et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. N Engl J Med 2004;351:1731-40.
3. Munn LL, Padera TP**.** Imaging the lymphatic system. Microvasc Res. 2014;0: 55–63.
4. Yankeelov TE, Pickens DR, Price RR**.** Quantitative MRI in Cancer: CRC Press 2011.
5. Bipat S, Glas AS, Slors FJM, et al. Rectal Cancer: Local Staging and Assessment of Lymph Node Involvement with Endoluminal US, CT, and MR Imaging. Radiology. 2004;232(3):773-83.
6. Jhaveri KS, Hosseini-Nik H**.** MRI of Rectal Cancer: An Overview and Update on Recent Advances. Am J Roentgenol 2015; 205(1):42-55.
7. Loftås P, Sturludóttir M, Hallböök O, et al. Assessment of remaining tumour involved lymph nodes with MRI in patients with complete luminal response after neoadjuvant treatment of rectal cancer. Br J Radiol. 2018;91(1087):20170938.
8. Yang X, Chen Y, Wen Z, et al. Role of Quantitative Dynamic Contrast-Enhanced MRI in Evaluating Regional Lymph Nodes With a Short-Axis Diameter of Less Than 5 mm in Rectal Cancer. Am J Roentgenol. 2019;212:77-83.
9. Santiago I, Santinha J, Ianus A, et al. Susceptibility perturbation MRI (SPI) maps tumor infiltration into mesorectal lymph nodes. Cancer Res. 2019:canres.3682.2018.
10. Herneth AM, Mayerhoefer M, Schernthaner R, et al. Diffusion weighted imaging: Lymph nodes. Radiology. 2010;76(3):398-406.
11. Koh DM, Collins D**.** Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. American Journal of Roentgenology. 2007;188(6):1622-35.
12. Reynaud O**.** Time-Dependent Diffusion MRI in Cancer: Tissue Modeling and Applications. Front. Phys. 2017;5:58.
13. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. . Neoplasia. 2009;11:102-25.

14. Surov A, Meyer HJ, Wienke A**.** Correlation between apparent diffusion coefficient (ADC) and cellularity is different in several tumors: a meta-analysis. Oncotarget. 2017;8(35):59492–9.
15. Sui WF, Chen X, Peng ZK, et al. The Diagnosis of Metastatic Axillary Lymph Nodes of Breast Cancer By Diffusion Weighted Imaging: a meta-analysis and systematic review. World J Surg Oncol. 2016;14:155.
16. Pekçevik Y, Çukurova I, Arslan IB**.** Apparent diffusion coefficient for discriminating metastatic lymph nodes in patients with squamous cell carcinoma of the head and neck. Diagn Interv Radiol. 2015;21(5):397–402.
17. Lim HK, Lee JH, Baek HJ, et al. Is diffusion-weighted MRI useful for differentiation of small non-necrotic cervical lymph nodes in patients with head and neck malignancies? Korean J Radiol. 2014;15(6):810-6.
18. Hasbahceci M, Akcakaya A, Memmi N, et al. Diffusion MRI on lymph node staging of gastric adenocarcinoma. Quant Imaging Med Surg. 2015;5(3): 392–400.
19. Cho EY, Kim SH, Yoon JH, et al. Apparent diffusion coefficient for discriminating metastatic from non-metastatic lymph nodes in primary rectal cancer. Eur J Radiol. 2013;82(11):e662-8.
20. Jensen JH, Helpern JA, Ramani A, et al. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med. 2005;53:1432-40.
21. Rosenkrantz AB, Padhani AR, Chenevert TL, et al. Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice. Journal of Magnetic Resonance Imaging. 2015;42(5):1190-202.
22. Shen L, Zhou G, Tang F, et al. MR diffusion kurtosis imaging for cancer diagnosis: A meta-analysis of the diagnostic accuracy of quantitative kurtosis value and diffusion coefficient. Clinical Imaging. 2018;52:44-56.
23. Jansen JFA, Stambuk HE, Koutcher JA, Shukla-Dave A**.** Non-Gaussian Analysis of Diffusion-Weighted MR Imaging in Head and Neck Squamous Cell Carcinoma: A Feasibility Study. American Journal of Neuroradiology. 2010;31(4):741-8.
24. Yamada I, Sakamoto J, Kobayashi D, et al. Diffusion kurtosis imaging of endometrial carcinoma: Correlation with histopathological findings. Magn Reson Imaging. 2019;57:337-46.
25. Yu J, Dai X, Zou HH, et al. Diffusion kurtosis imaging in identifying the malignancy of lymph nodes during the primary staging of rectal cancer. Colorectal Dis. 2018;20(2):116-25.
26. Le Bihan D, Breton E, Lallemand D, et al. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. Radiology. 1988;168.
27. Iima M, Le Bihan D**.** Clinical Intravoxel Incoherent Motion and Diffusion MR Imaging: Past, Present, and Future. Radiology. 2016;278(1):13-32.
28. Qiu L, Liu X, Liu S, et al. Role of quantitative intravoxel incoherent motion parameters in the preoperative diagnosis of nodal metastasis in patients with rectal carcinoma. J. Magn. Reson. Imaging. 2016;44: 1031-9.
29. Yu XP, Wen L, Hou J, et al. Discrimination between Metastatic and Nonmetastatic Mesorectal Lymph Nodes in Rectal Cancer Using Intravoxel Incoherent Motion Diffusion-weighted Magnetic Resonance Imaging. Acad Radiol. 2016;23:479–85.

30. Wu WC, Yang SC, Chen YF, et al. Simultaneous assessment of cerebral blood volume and diffusion heterogeneity using hybrid IVIM and DK MR imaging: initial experience with brain tumors. Eur Radiol. 2017;27(1):306-14.
31. Fujima N, Sakashita T, Homma A, et al. Utility of a Hybrid IVIM-DKI Model to Predict the Development of Distant Metastasis in Head and Neck Squamous Cell Carcinoma Patients. Magn Reson Med Sci. 2018;17(1):21-7.
32. Wang YJ, Xu XQ, 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. 2018;59(6):672-80.
33. Panagiotaki E, Walker-Samuel S, Siow B, et al. Noninvasive quantification of solid tumor microstructure using VERDICT MRI. Cancer Res. 2014;74:1902-12.
34. White NS, McDonald CR, Farid N, et al. Diffusion-weighted imaging in cancer: physical foundations and applications of restriction spectrum imaging. Cancer Res. 2014;74:4638-52.
35. Reynaud O, Winters KV, Hoang DM, et al. Surface-to-volume ratio mapping of tumor microstructure using oscillating gradient diffusion weighted imaging. Magn Reson Med. 2017;76(1):237–47.
36. Jiang X, Li H, Xie J, et al. In vivo imaging of cancer cell size and cellularity using temporal diffusion spectroscopy. Magn Reson Med. 2017;78(1):156–64.
37. Liang S, Panagiotaki E, Bongers A, et al. Information-based ranking of 10 compartment models of diffusion-weighted signal attenuation in fixed prostate tissue. NMR Biomed. 2016;29:660-71.
38. Szczepankiewicz F, van Westen D, Englund E, et al. The link between diffusion MRI and tumor heterogeneity: Mapping cell eccentricity and density by diffusional variance decomposition (DIVIDE). NeuroImage. 2016;In Press.
39. Liang C-Y, Chen M-D, Zhao X-X, et al. Multiple mathematical models of diffusion-weighted magnetic resonance imaging combined with prognostic factors for assessing the response to neoadjuvant chemotherapy and radiation therapy in locally advanced rectal cancer. European Journal of Radiology. 2019;110:249-55.
40. Xu J, Jiang X, Li H, et al. Magnetic resonance imaging of mean cell size in human breast tumors. ArXiv preprint arXiv:1905.07818. 2019.
41. Bonet-Carne E, Tariq M, Pye H, et al. Histological Validation of in-vivo VERDICT MRI for Prostate using 3D Personalised Moulds. Proceedings of the Joint Annual Meeting ISMRM-ESMRMB 2018, 2018. Paris.
42. Bailey C, Bourne RM, Siow B, et al. VERDICT MRI validation in fresh and fixed prostate specimens using patient-specific moulds for histological and MR alignment. NMR Biomed. 2019;32(5):e4073.
43. Panagiotaki E, Chan R, Dikaios N, et al. Microstructural characterisation of normal and malignant human prostate tissue with VERDICT MRI. Investigative Radiology. 2014;50(4):218-22.
44. Johnston EWJ, Bonet-Carne E, Ferizi U, et al. VERDICT MRI for Prostate Cancer: Intracellular Volume Fraction versus Apparent Diffusion Coefficient. Radiology. 2019;291:391–7.
45. Zaccagna F, Riemer F, Priest AN, et al. Non-invasive assessment of glioma microstructure using VERDICT MRI: correlation with histology. Eur J Radiol. 2019.

46. Reynaud O, Winters KV, Hoang DM, et al. Pulsed and oscillating gradient MRI for assessment of cell size and extracellular space (POMACE) in mouse gliomas. NMR Biomed. 2016;29(10): 1350-63.
47. Jiang X, Li H, Xie J, et al. Quantification of cell size using temporal diffusion spectroscopy. . Magnetic Resonance in Medicine. 2016;75:1076-85.
48. Sahalan M**.** Diffusion-weighted Imaging of Lymph Node Tissue. University of Sydney; 2019.
49. Panagiotaki E, Schneider T, Siow B, et al. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. NeuroImage. 2012;59:2241-54.
50. Li H, Jiang X, Xie J, et al. Impact of transcytolemmal water exchange on estimates of tissue microstructural properties derived from diffusion MRI. Magn Reson Med. 2017;77(6):2239-49.
51. Alexander DC, Dyrby TB**.** Diffusion imaging with stimulated echoes: signal models and experiment design. arXiv:1305.7367 2013.
52. Neuman CH**.** Spin echo of spins diffusing in a bounded medium. J. Chem. Phys. 1974;60:4508.
53. Merboldt KD, Hanicke W, Frahm J**.** Self-diffusion NMR imaging using stimulated echoes. Journal of Magnetic Resonance (1969). 1985;64(3):479-86.
54. Teruel JR, Cho GY, Moccaldi RM, et al. Stimulated echo diffusion tensor imaging (STEAM-DTI) with varying diffusion times as a probe of breast tissue. J. Magn. Reson. Imaging. 2017;45(1):84-93.
55. Nolte UG, Finsterbusch J, Frahm J**.** Rapid isotropic diffusion mapping without susceptibility artifacts: Whole brain studies using diffusion-weighted single-shot STEAM MR imaging. Magn Reson Med. 2000;44(5):731-6.
56. DeLong ER, DeLong DM, Clarke-Pearson DL**.** Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 1988;44(3):837-45.
57. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for clinical management of rectal cancer: Updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. Eur Radiol. 2018;28(4):1465-75.
58. Fan X, Macleod K, Mustafi D, et al. Correlation of In Vivo and Ex Vivo ADC and T2 of In Situ and Invasive Murine Mammary Cancers. PLoS One. 2015;10(7):e0129212.
59. Korteweg MA, Zwanenburg JJM, van Diest PJ, et al. Characterization of ex vivo healthy human axillary lymph nodes with high resolution 7 Tesla MRI. Eur Radiol. 2011;21(2):310–7.
60. Giannakidis A, Gullberg GT, Pennell DJ, Firmin DN**.** Value of Formalin Fixation for the Prolonged Preservation of Rodent Myocardial Microanatomical Organization: Evidence by MR Diffusion Tensor Imaging. Anat Rec (Hoboken). 2016;299(7):878-87.
61. Shatil AS, Uddin MN, Matsuda KM, Figley CR**.** Quantitative Ex Vivo MRI Changes due to Progressive Formalin Fixation in Whole Human Brain Specimens: Longitudinal Characterization of Diffusion, Relaxometry, and Myelin Water Fraction Measurements at 3T. Front Med (Lausanne). 2018;5:31.
62. Pfeuffer J, Flögel U, Dreher W, Leibfritz D**.** Restricted diffusion and exchange of intracellular water: theoretical modelling and diffusion time dependence of 1H NMR measurements on perfused glial cells. NMR Biomed. 1998;11(1):19-31.

63. Tian X, Li H, Jiang X, et al. Evaluation and comparison of diffusion MR methods for measuring apparent transcytolemmal water exchange rate constant. Journal of Magnetic Resonance. 2017;275: 29–37.

64. Lee HH, Fieremans E, Novikov DS**.** What dominates the time dependence of diffusion transverse to axons: Intra- or extra-axonal water? NeuroImage. 2018;182:500-10.

Tables:

a) Parameter	Pre-clinical scan parameters	
Echo time (ms)	6.5	
Repetition time (ms)	2800	
Slice Thickness (mm)	0.7	
In-plane resolution (mm^2)	0.14 x 0.14	
Matrix size	70 x 70	
Bandwidth (kHz)	34.722	
Gradient duration (ms)	1	
# Gradient directions	6	
B values (s/mm^2)	0, 500, 1000	0, 1500, 2000
Signal averages	1	2
Diffusion time (ms)	5, 10, 20, 40, 70, 100, 150	10, 20, 30, 50
Mixing time (ms)	2.2, 7.2, 17.2, 37.2, 67.2, 97.2, 147,2	7.2, 17.2, 27.2, 47.2

b)	Clinical scan parameters		
Parameter	**Small b values**	**Medium b values**	**High b values**
Echo time (ms)	65	65	65
Repetition time (ms)	3550	3755	9561
Flip angle	90	90	90
Slice thickness (mm)	5	5	5
Gap (mm)	0	0	0
Matrix	128x126	128x126	128x126
Field-of-view (mm)	320x320	320x320	320x320
In-plane resolution (mm^2)	2.5x2.5	2.5x2.5	2.5x2.5
Signal averages	1	1	2
EPI factor	63	63	63
Bandwidth (kHz)	29.1	21.7	21.7
dB/dt (T/s)	91.5	41.2	41.2
Acquisition Duration	2min26s	2min34s	15min09s
Fat suppression	SPAIR	SPAIR	SPAIR
Halfscan factor	0.696	0.696	0.667
B values	0, 50, 100, 200	0, 500, 1000	0, 1500, 2000, 2500
Gradient duration (ms)	15	16.6	16.9
Diffusion time (ms)	35.7	32.5	32.8

Table 1. Imaging parameters for a) pre-clinical dMRI acquisition at 16.4T with stimulated echo preparation and line-by-line readout and b) clinical dMRI acquisition at 1.5T with spin-echo preparation and multi-shot EPI readout.

	ADC	IVIM			Kurtosis		IVIM - Kurtosis			
	D $(\mathrm{ms}/\mu\mathrm{m}^2)$	f	D $(\mathrm{ms}/\mu\mathrm{m}^2)$	D* $(\mathrm{ms}/\mu\mathrm{m}^2)$	D $(\mathrm{ms}/\mu\mathrm{m}^2)$	k	**D $(\mathrm{ms}/\mu\mathrm{m}^2)$**	f	D* $(\mathrm{ms}/\mu\mathrm{m}^2)$	k
Benign	0.96 ± 0.34	0.16 ± 0.12	0.70 ± 0.28	12.9 ± 10.9	1.12 ± 0.39	1.14 ± 0.34	**0.80 ± 0.30**	0.15 ± 0.13	13.7 ± 11	1.09 ± 0.62
Malignant	1.08 ± 0.36	0.13 ± 0.15	0.85 ± 0.39	8.0 ± 7.5	1.27 ± 0.42	0.94 ± 0.27	**1.02 ± 0.41**	0.11 ± 0.13	10.1 ± 8.2	0.88 ± 0.49
p-value	0.26	0.66	0.11	0.074	0.22	0.12	**0.030**	0.32	0.24	0.32

Table 2 Parameter values for benign and malignant nodes given by the ADC, IVIM, Kurtosis and IVIM-Kurtosis models applied to clinical data.

Figures:

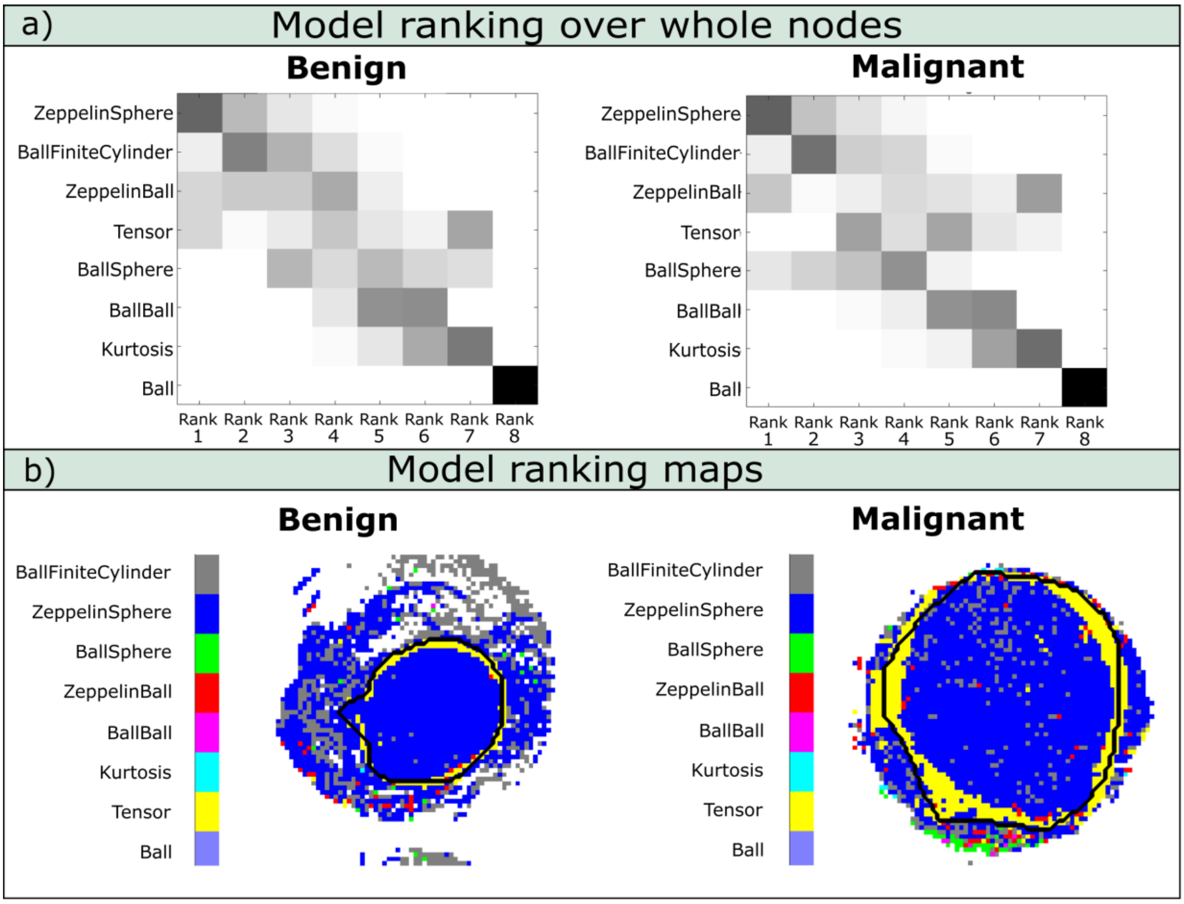

Figure 1 a) Model ranking for benign (left) and malignant (right) lymph nodes. The models are fitted voxelwise and the grey scale reflects the frequency with which a model occupies a certain rank. Thus, the model on top provides the best fit in most of voxels, while the model on the bottom provides the worst fit in most of voxels. The models are roughly ordered according to how many times they provide a certain rank, from best (top) to worst (bottom). b) Maps showing the best fitting model in each voxel of a benign (left) and malignant (right) node.

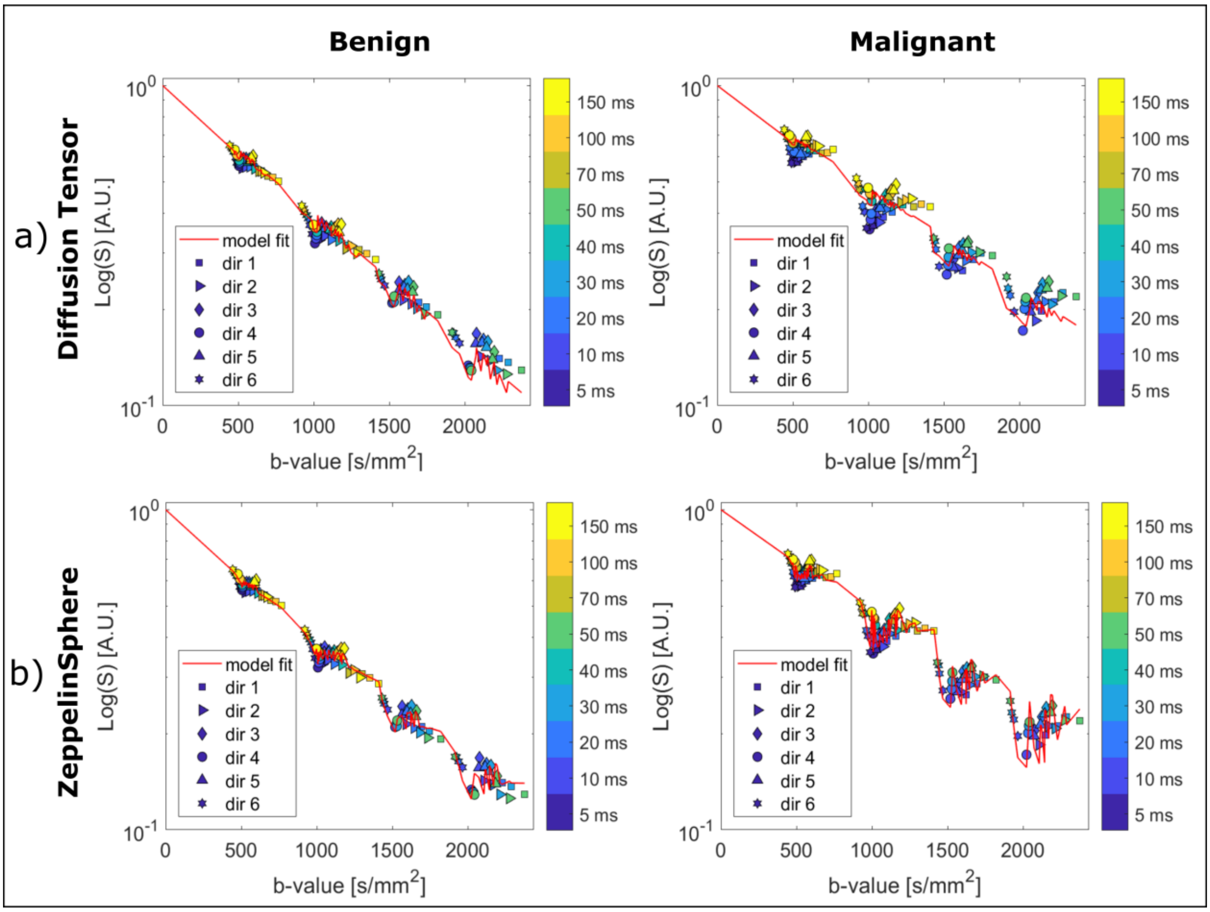

Figure 2 Example of voxelwise model fitting for a) Tensor and b) ZeppelinSphere models in a benign and a malignant node. The different symbols show different gradient directions, and the colour encodes different diffusion times. The red line shows the model fit to the data.

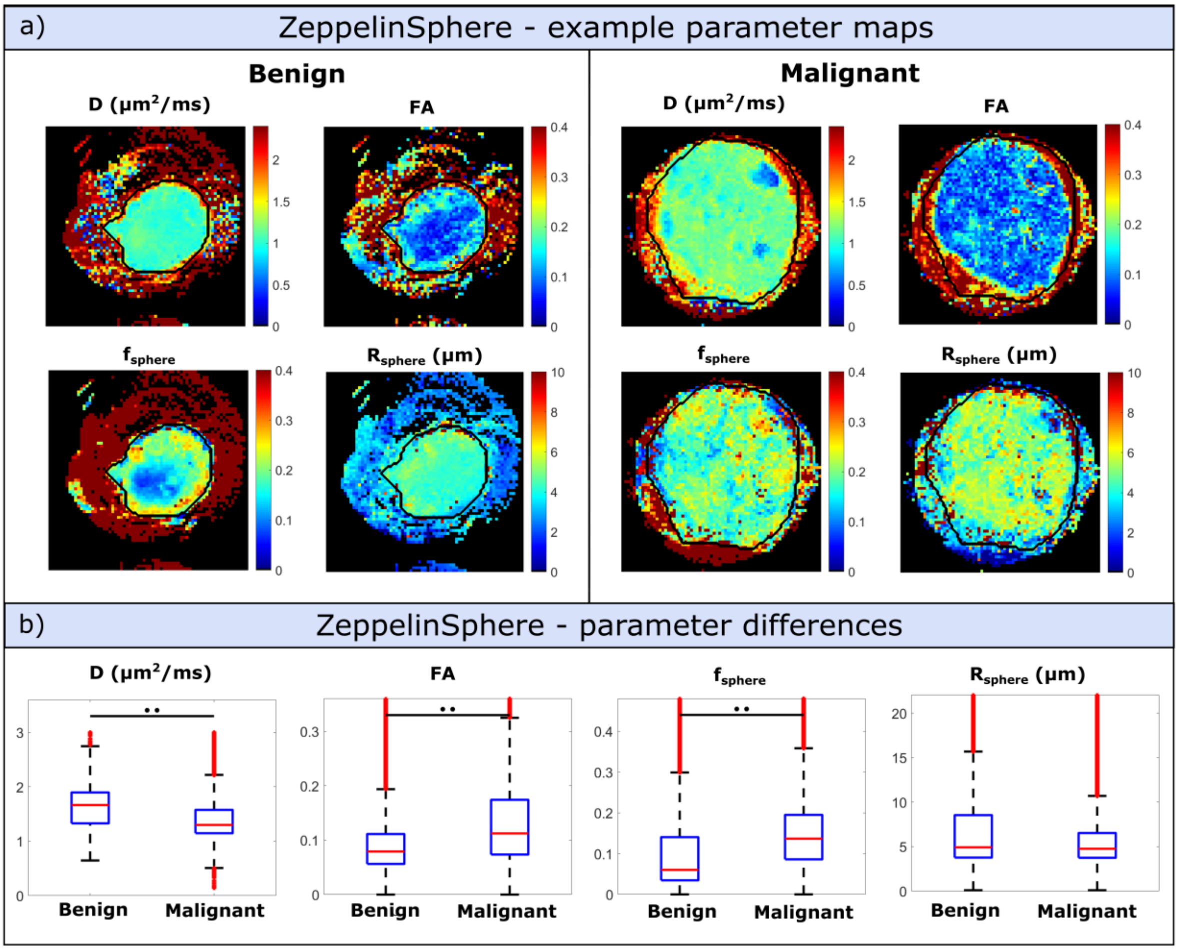

Figure 3 a) Parameter maps derived from the ZeppelinSphere model (Di, FA, f_{sphere} and R_{sphere}) for a benign and malignant lymph node. b) Di, FA, f_{sphere} and R_{sphere} comparison for benign and malignant nodes.

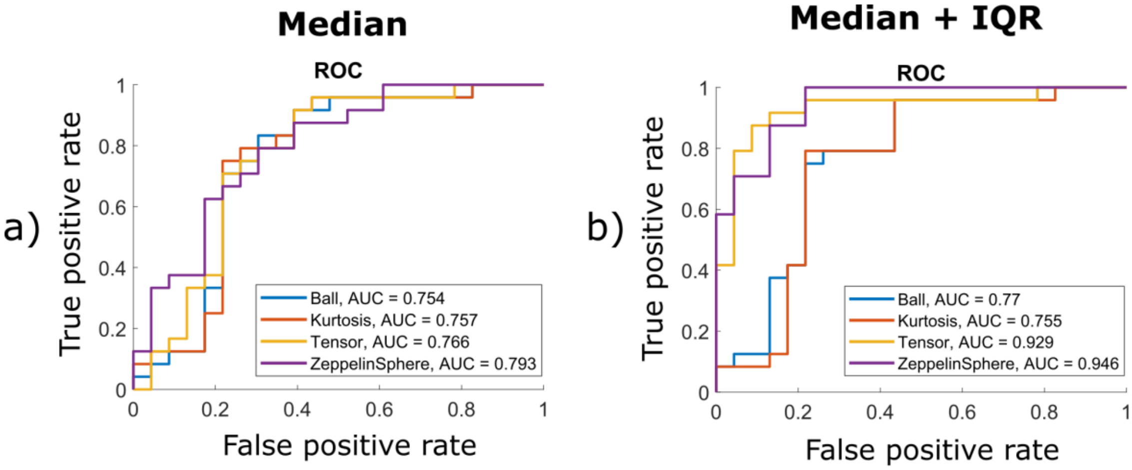

Figure 4 ROC curves for the Ball, Kurtosis, Tensor and ZeppelinSphere model when including for each model parameter the a) median values and b) median values and IQR.

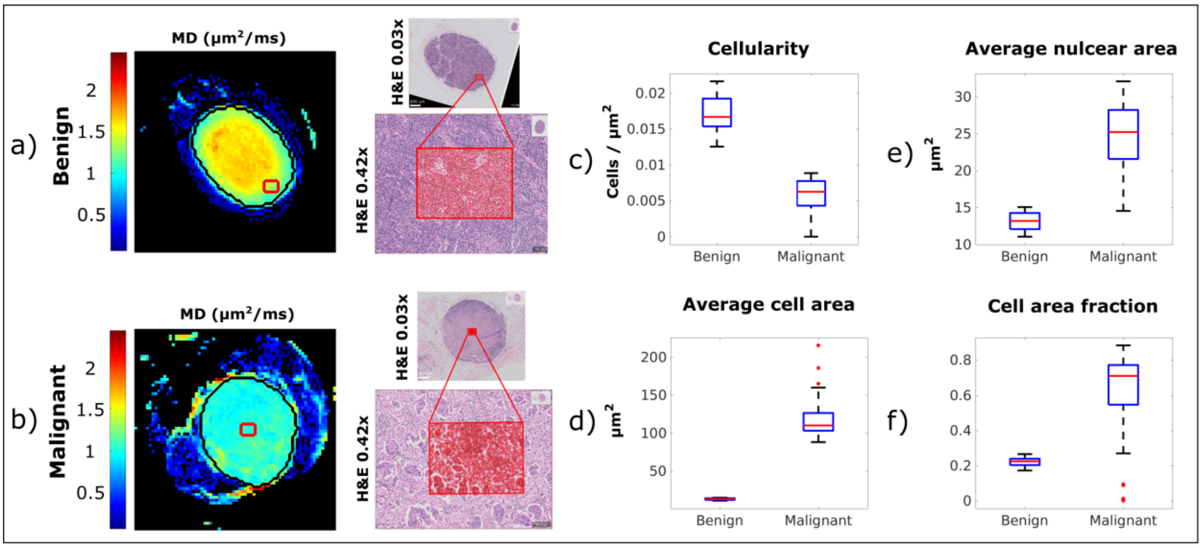

Figure 5 Example of corresponding ROIs between MRI and histology for a) benign node and b) malignant nodes. Comparison of histology features between benign and malignant nodes: c) cellularity, d) average cell area, e) average nuclear area and f) cell area fraction.

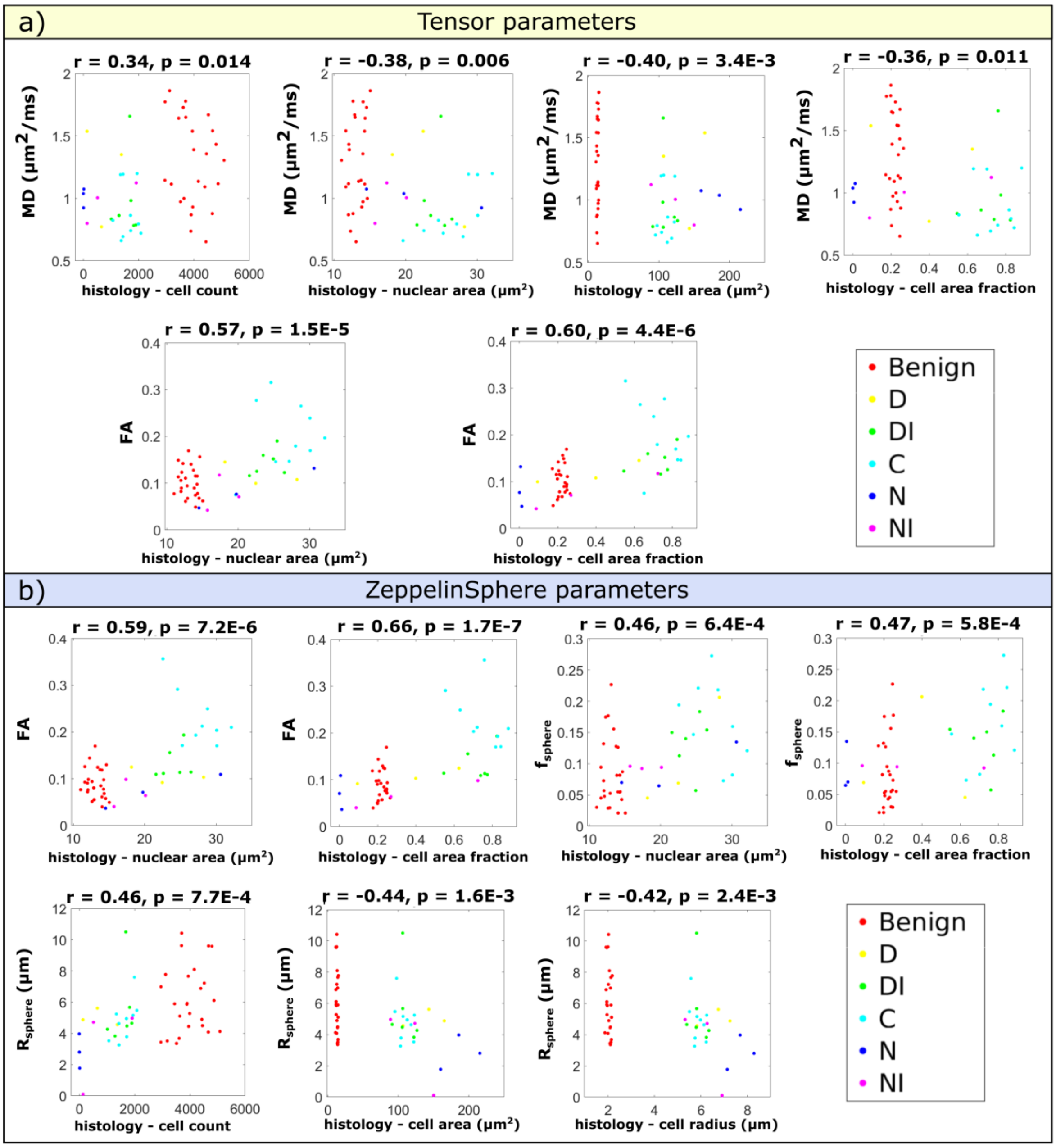

Figure 6 Correlation between MRI derived model parameters and histology features for the a) Tensor model and b) ZeppelinSphere model. Only parameters with significant correlations with r> 0.3 are depicted. For the R_{sphere} parameter we also show the correlation with the estimated cell radius from histology, which is calculated from the average cell area assuming a circular shape. Different marker colours represent different node types. Benign nodes are depicted in red, while malignant nodes are depicted in different colours depending on their malignancy pattern (D – desmoplastic; DI – desmoplastic with inflammation; C - cellular N – necrotic; NI – necrotic with inflammation). The following correlation strengths are assumed based on the values of the correlation coefficient r: <0.3 = negligible; 0.3 to 0.5 = weak; 0.5 to 0.7 = moderate; 0.7 to 0.9 = strong; 0.9 to 1.00 = very strong (63).

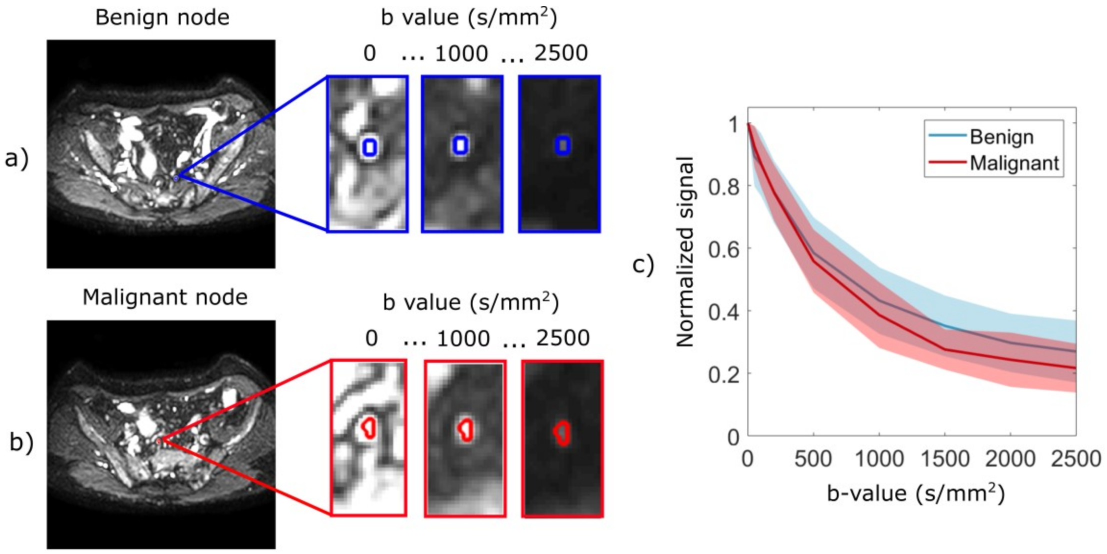

Figure 7 a)-b) Diffusion weighted images of a benign and a malignant lymph node, respectively, with full FOV b0 images and details of the lymph node ROI delineation for three b-values (0, 1000 and 2500 s/mm^2). c) Average signal decay for benign and malignant lymph nodes as a function of b-value.

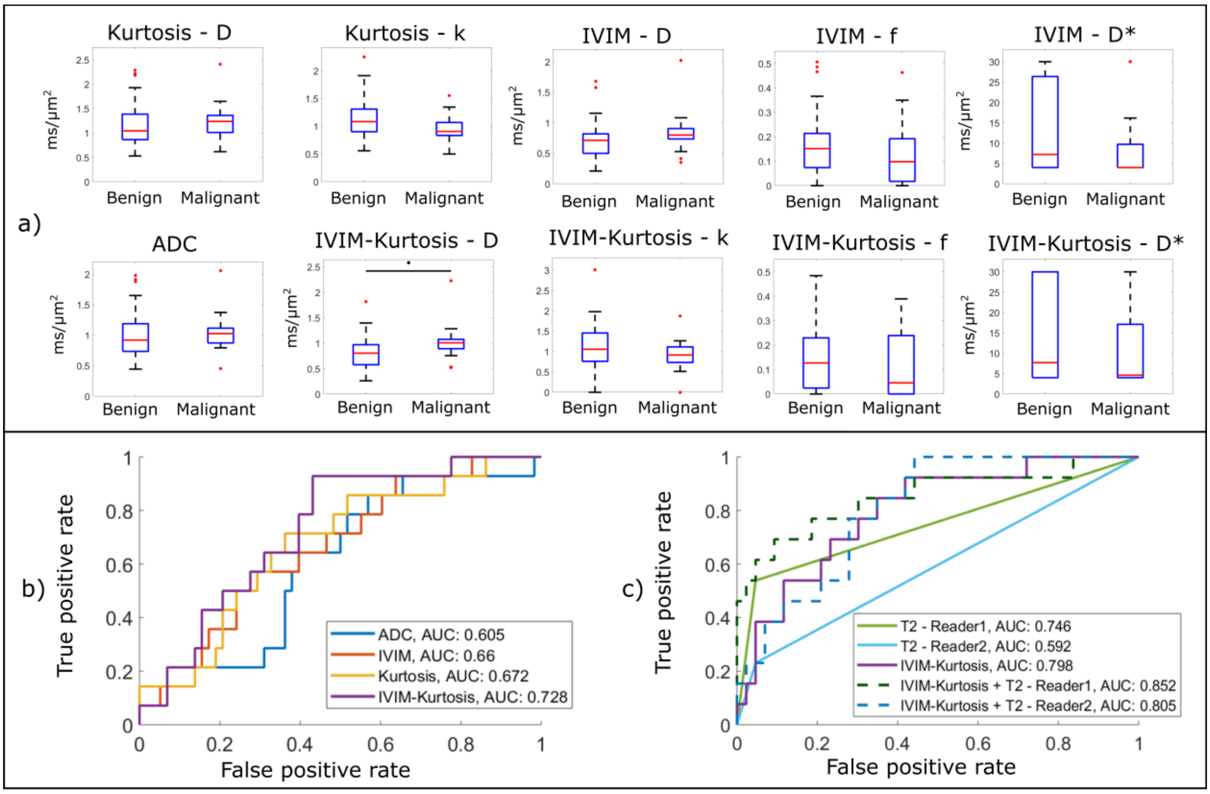

Figure 8 a) Estimated parameter values of the various diffusion model for benign and malignant lymph nodes (Kurtosis and IVIM on the top row; ADC and IVIM-Kurtosis on the bottom row). b) ROC curves for the ADC, IVIM, Kurtosis and IVIM-Kurtosis models. c) ROC

curves for the IVIM-Kurtosis model, as well as for the node classification provided by two different radiologists. This ROC curve is evaluated on a subset of 56 nodes.

Supporting Information Table S1 Description of the diffusion models fitted to the pre-clinical ex-vivo data including model name, equation, number of fitted parameters and the range of parameter values.