Article:
Skaarup, M, Lundemann, MJ, Darkner, S et al. (8 more authors) (2021) A framework for voxel-based assessment of biological effect after proton radiotherapy in pediatric brain cancer patients using multi-modal imaging. Medical Physics, 48 (7). pp. 4110-4121. ISSN 0094-2405

https://doi.org/10.1002/mp.14989

A framework for voxel-based assessment of biological effect after proton radiotherapy in pediatric brain cancer patients using multi-modal imaging

Framework for in-vivo biological effect

Mikkel Skaarup[1,2]*, Michael Juncker Lundemann[1], Sune Darkner[3], Morten Jørgensen[1], Lisbeth Marner [4,5], Dragan Mirkovic[6], David Grosshans[6], Christopher Peeler[6], Radhe Mohan[6], Ivan Richter Vogelius[1,7], Ane Appelt[8]

[1]Department of Oncology, Rigshospitalet, Copenhagen, Denmark; [2]Faculty of Science, Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark; [3]Department of Computer Science, University of Copenhagen, Copenhagen, Denmark; [4]Department of Clinical Physiology, Nuclear Medicine and PET, Rigshospitalet, Copenhagen, Denmark; [5]Department of Clinical Physiology and Nuclear Medicine, Copenhagen University Hospital Bispebjerg, Denmark; [6]Department of Radiation Physics, Division of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX; [7]Faculty of Health and Medical Science, Copenhagen University, Copenhagen, Denmark; [8]Leeds Institute of Medical Research at St James's, University of Leeds, and Leeds Cancer Centre, St. James's University Hospital, Leeds, United Kingdom

*Corresponding author: Mikkel Skaarup, Department of Oncology, Rigshospitalet, Blegdamsvej 9, 2100 København Ø, Denmark. (+45) 35458931. mikkel.skaarup@regionh.dk

ORCID ID - Mikkel Skaarup: https://orcid.org/0000-0002-3460-9741

A framework for voxel-based assessment of biological effect after proton radiotherapy in pediatric brain cancer patients using multimodal imaging

Abstract

Introduction

The exact dependence of biological effect on dose and linear energy transfer (LET) in human tissue when delivering proton therapy is unknown. In this study we propose a framework for measuring this dependency using multi-modal image-based assays with deformable registrations within imaging sessions and across time.

Materials and Methods

3T MRI scans were prospectively collected from 6 pediatric brain cancer patients before they underwent proton therapy treatment, and every 3 months for a year after treatment. Scans included T1-weighted with contrast enhancement (T1), T2-FLAIR (T2) and fractional anisotropy (FA) images. In addition, the planning CT, dose distributions and Monte Carlo-calculated LET distributions were collected.

A multi-modal deformable image registration framework was used to create a dataset of dose, LET and imaging intensities at baseline and follow-up on a voxel-by-voxel basis. We modelled the biological effect of dose and LET from proton therapy using imaging changes over time as a surrogate for biological effect.

We investigated various models to show the feasibility of the framework to model imaging changes. To account for inter- and intra-patient variations we used a nested generalized linear mixed regression model. The models were applied to predict imaging changes over time as a function of dose and LET for each modality.

Results

Using the nested models to predict imaging changes, we saw a decrease in the FA signal as a function of dose; however, the signal increased with increasing LET. Similarly, we saw an increase in T2 signal as a function of dose, but a decrease in signal with LET. We saw no changes in T1 voxel values as a function of either dose or LET.

Conclusions

The imaging changes could successfully model biological effect as a function of dose and LET using our proposed framework[1]. Due to the low number of patients, the imaging changes observed for FA and T2 scans were not marked enough to draw any firm conclusions.

Keywords: Multi-modality registration; Diffusion MRI; Radiobiology of protons; Proton therapy; MRI response assessment

Introduction

Proton therapy takes advantage of the finite range of protons, enabling sparing of normal tissue at the distal end of the treatment beam[2,3]. There are, however, a number of geometric uncertainties related to treatment delivery, such as range- and setup-uncertainties[4–7], which may limit the clinical benefit. In addition, there are indications that the clinical standard assumption of a constant relative biological effectiveness (RBE) across the beam is not accurate, but rather depends on the local linear energy transfer (LET)[8], as has been shown in in-vitro studies[9–11]. More recently, clinical observations have supported this conjecture, for example using Monte Carlo simulation[12], retrospectively evaluating radiographic abnormalities in the chest wall[13], and modelling post-treatment image changes as delineated by a physician[14,15]. The available data suggest that the biological effect of protons increases as the LET increase, typically towards the end of the beam. This may negate part of the beneficial effect of proton therapy unless properly accounted for in planning.

Various models have been suggested to describe a variable biological effect[9,10,16]. However, these models are based on in-vitro data, and the clinical data to validate them are missing. To transfer our understanding of variable biological effects to clinical treatment plan optimization, we need better estimates of the relationship between local effect, dose and LET, based on in vivo data. This includes dependence on tissue type, fractionation, alpha/beta ratio and other effects. To accurately model this, a framework for in-vivo assessment of local biological effect is needed.

Imaging has been used as a marker for radiation damage in a number of settings, for example to assess treatment response[17,18], outcome predictions[19] or for adaptive treatments[20,21]. These studies use a variety of MRI, PET and CT images, and have been able to relate a per-voxel radiation dose to a local response. As LET has varying localized effects in-vivo and because of inhomogeneous dose distributions, using imaging to investigate a variable biological effect of protons is an ideal choice. Fractional anisotropy (FA) imaging maps myelinated white brain matter, and it has been shown that FA maps loose signal due to radiation damage[22–

[24] and that FA changes are associated with cognitive impair after trauma[25]. A link between damage of white matter, following combined surgery and proton radiation and measured with FA maps, and cognitive performance has also been shown[26]. It is plausible that the FA maps can serve to measure relevant components of the biological damage of radiation.

We therefore aimed to develop a multimodal imaging framework for modelling imaging changes over time on a voxel-by-voxel level, thus being able to link imaging changes (as a surrogate for biological effectiveness) to the dose and LET distribution. The framework takes inter- and intra-variation between patients into account by using a sophisticated deformable registration strategy between template timepoints, where a template consists of different relevant imaging modalities acquired at the same timepoint. Our proposed method supports multimodal imaging, which we demonstrate using different magnetic resonance imaging (MRI) sequences and CT images.

Materials and Methods

In brief, we enrolled patients in a prospective study to longitudinally assess normal tissue imaging response after proton therapy. A complex series of rigid and deformable registrations provided us with a voxel-by-voxel data set. Finally, three models of LET and dose dependence were chosen to show the ability of the imaging framework to describe imaging changes over time. We chose to use T1, T2 and FA MRI images in this study.

Patient cohort and image acquisition

For image-based quantification of local effect as a function of dose and LET, a robust data acquisition protocol is required. Pediatric brain cancer patients from the Capital Region of Denmark were prospectively included in a clinical study, approved by the regional Scientific Ethics Committee (H-3-2014-135). They were enrolled prior to referral for proton therapy, which was delivered at the MD Anderson Cancer Center, University of Texas, Houston, USA. Radiotherapy and chemotherapy protocols were in line with international standards for each cancer type[27,28]. All patients were treated with passive scattering proton

therapy. Initial assessment, adjuvant treatment, and follow-up were conducted at Rigshospitalet, Denmark. After informed consent by the patient or the patient's guardian, we acquired dedicated study MRI scans of the patients before and after their proton therapy, alongside standard clinical imaging. The scans were acquired two weeks before proton therapy was initiated (baseline), and at follow-ups two months after proton therapy completion and then at three-monthly intervals for a year. The dedicated study MR protocol included the following standard clinical structural sequences: a T1 weighted magnetization-prepared rapid acquisition with gradient echo (MPRAGE, 1mm x 1mm x 1mm voxel size) with contrast agent (gadolinium), and an axial T2 and coronal Fluid-attenuated inversion recovery (Flair, 0.7188mm x 0.7188mm x 6mm voxel size). In the remainder of this paper, we will refer to these two sequences simply as "T1" and "T2". Additional clinical sequences varied on a patient specific basis, for example T2 dark fluid sequences were included where appropriate. Further, we performed study protocol-specific MR diffusion tensor imaging (DTI) designed to measure myelinated white brain matter. The latter was made by measuring directional diffusion in thirty main directions. All other DTI parameters were equal with b value = 1,000 $mm^{-2,}$ repetition time/echo time = 4900/ 90ms, field of view = 230mm x 230mm x 190mm, acquisition matrix = 128x128x37 voxels, flip angle = 90°. The three-dimensional directional diffusion images were used to create a FA map[18]. MRI scans were performed on a 3T Siemens Prisma Scanner (Siemens, Erlangen, Germany) using a 64-channel head coil or at a 3T Siemens Biograph mMR (Siemens, Erlangen, Germany) with a similar protocol. Corrections for motion and eddy currents of the FA map were done using FSL python library for DTI brain imaging data**.** As part of the proton treatment planning, all patients received a CT scan for dose calculations which was also included in the study. The planning CT scan was used for dose planning and LET calculation. One patient (patient 2) was also enrolled in a study of [^{18}F]FET PET/MRI examinations of patients with pediatric CNS tumors[29] approved by the regional Scientific Ethics Committee (H-6-2014-095) and had two of his MRI scans performed at the PET/MRI scanner to reduce the number of scanning procedures. Example imaging and treatment information for one patient can be seen in Fig. 1.

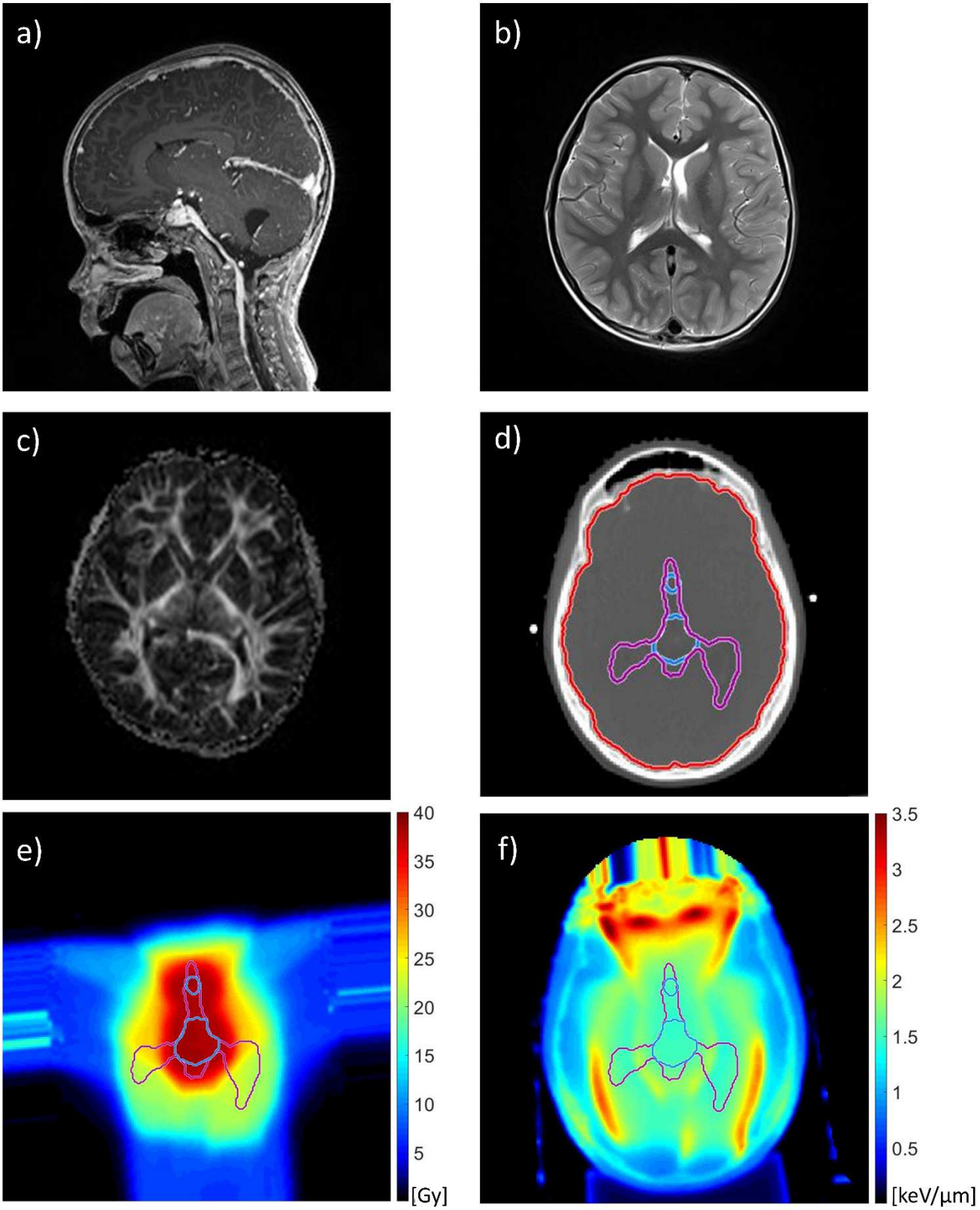

a)
b)
c)
d)
e)
40
35
30
25
20
15
10
5
[Gy]
f)
3.5
3
2.5
2
1.5
1
0.5
[keV/µm]

Figure 1. An example of the scans acquired at each timepoint (the CT, dose and linear energy transfer (LET) is only acquired at baseline) for patient 6. a) sagittal T1-weighted MRI, b) T2-weighted MRI, c) fractional anisotropy (FA) map, d) baseline CT with corresponding brain, CTV receiving 24 Gy (purple) and CTV boost (of 16 Gy, in blue) delineations, e) dose distribution and f) LET for that plan.

Dose and LET calculation

All treatment plans were recalculated using Monte Carlo in order to produce dose and LET distributions. Clinical treatment plans were extracted from the Eclipse treatment planning system (v9.0) in DICOM format and then used as input for an in-house Monte Carlo system. The system, called MC^2, uses the DICOM files to generate geometry specifications and other files describing the beam arrangement and delivery parameters. The files are then used as input for simulation with the Monte Carlo code MCNPX, as has been described previously [30,31]. The Monte Carlo code used a simulated number of particles that ensured an acceptable uncertainty (+/- 2%) of the dose distribution.

Physical doses per beam were summed to create a single dose distribution per patient and per voxel (for all beams in multi-phase plans). Similarly, per-beam LET maps were summed using the per-beam dose map as weighting factors, to create a single dose-weighted LET map per patient. Across the included patients the average dose per fraction was 1.73 Gy (1.6-1.8 Gy). Opposing beams or multiple beams equally spread out, with angles chosen individually for each patient, were used.

Image registration

The framework for image registration between time-points was a combination of rigid and deformable image registrations, designed to minimize the impact of patient positioning variations and anatomy changes over time and to be able to include multiple scans as required to assess the biological effect of radiation on a voxel-by-voxel basis. We used a previously published multi-modality image registration algorithm[32,33]. At each timepoint, separately for each patient, all scans acquired were rigidly registered within each timepoint to create a timepoint "template" of all included modalities. The rigid registration used the scan with the highest resolution in each template as the reference, which in our study was the T2

MRI. Between each follow-up and corresponding baseline templates, a deformable image registration was carried out. This registration was iterative between all scans in each template, i.e. each individual scan in the moving template was deformably registered to all the scans of the fixed template, iteratively updating the combined template registration matrix for each combination. This process was repeated until convergence of the combined template registration matrix was reached. As the CT scan was included in the baseline template for each patient, dose and LET were thus also mapped for each voxel. The result was a dataset of imaging intensities relative to baseline, dose and LET on a voxel-by-voxel basis for each follow-up. The voxel-size in the templates was 1x1x1 mm^3. An overview of the registration process can be seen in Fig. 2.

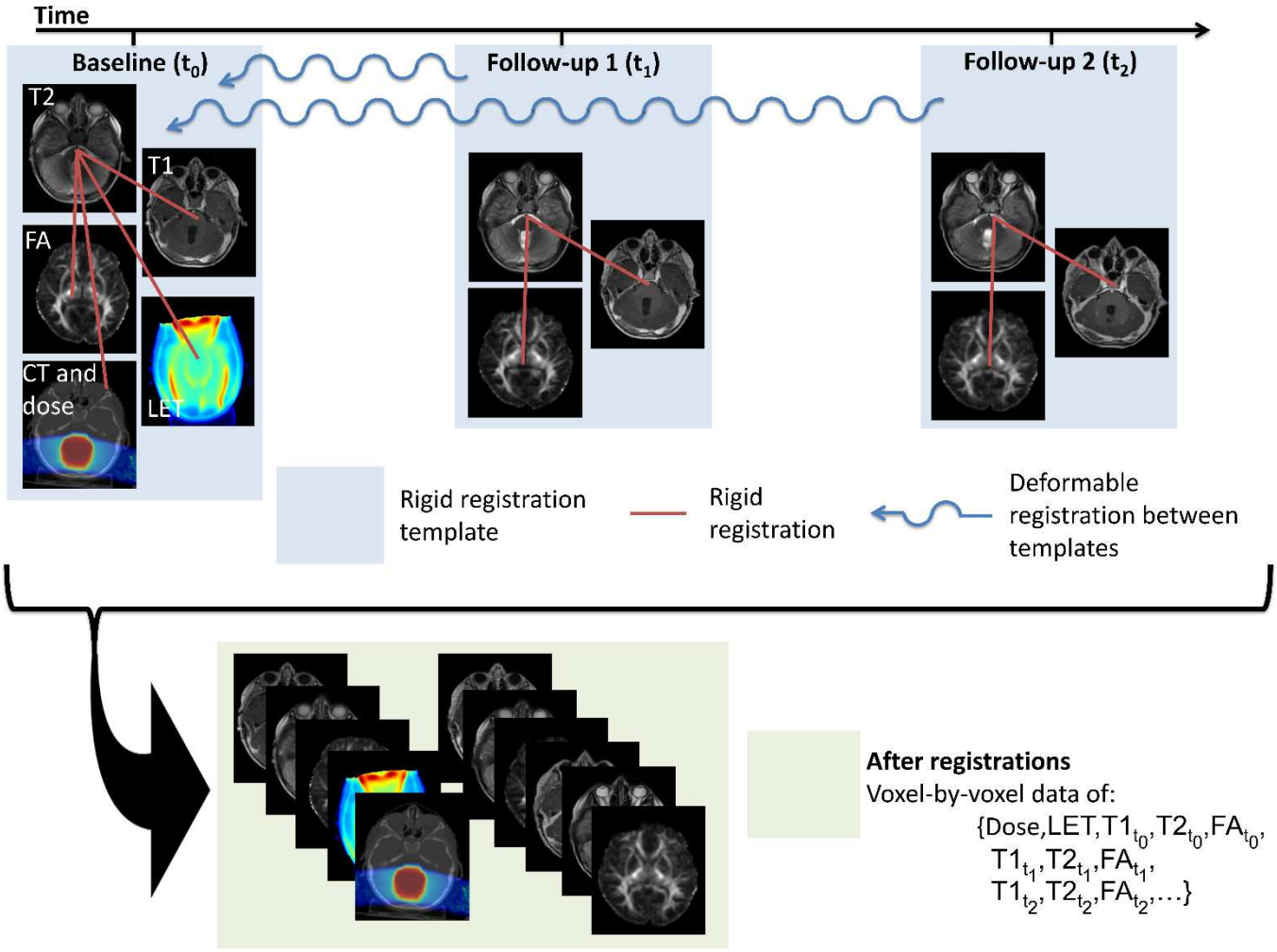


Figure 2. An overview of the registration process. The blue boxes represent scan sessions, called a template, taken at baseline and follow-ups at three-month intervals after completed radiotherapy. In each template the scans are rigidly registered to the T2-weighted MRI scan. Between each template, a deformable registration is carried out using information from all scans in both templates. This process leads to a voxel-by-voxel dataset of imaging changes since baseline of T1, T2 and fractional anisotropy (FA) MRI scan values as well as the dose and linear energy transfer (LET) of those voxels.

Data processing

Processing of the imaging signal was done to reduce the amount of noise in each scan and to focus on relevant regions of interest for data analysis. Two masks were created, one for FA MRI scans and one for T1/T2 weighted MRI scans. Both masks were initially restricted to cover only the brain as outlined on the CT for treatment planning (see Fig. 1), and further requirements were specified for each of the masks. In the

FA image, we isolated voxels where the baseline image showed a white matter tract signal, i.e. where the FA scan value was above the noise level. Specifically, we used a gaussian fit to the lower values of the voxel value frequency histogram for baseline FA images on a per-patient basis to filter out noise (see Supplementary Figure 1). A two-sigma cut-off was chosen for noise removal in the FA mask. The FA mask was then restricted to only include voxels with a baseline FA value above this cut-off (described in Supplementary Figure 1) and additionally restricted to voxels receiving doses above 35 Gy. This latter criteria was used as it has previously been shown that low doses do not appear to elicit a dose-response signal on FA imaging[24]. An additional explorative analysis used varying dose thresholds (5 and 20 Gy). The T1/T2 mask had the same requirement of 35 Gy, but no further signal restrictions. Masks were created based on the baseline scans but applied to follow-up data after registrations.

Data analysis

We analyzed the data using nested generalized linear mixed regression models to relate dose and LET to variation in image signal. This was done separately for the three MRI modalities, T1, T2 FLAIR and FA images. The general regression model was[34]:

$$\text{Follow-up}_{i,j,n} = a \cdot \text{Baseline}_{i,n} + b \cdot \text{f}\left(\text{dose}_{i,n}\right) + c \cdot \text{LET}_{i,n} \cdot \text{f}\left(\text{dose}_{i,n}\right) + \alpha_{i,j} \qquad (1)$$

where the *i*-subscript refers to an individual patient, while the *j*-subscript refers to the different follow-up sessions for that patient, and the *n*-subscript refers to a specific voxel for the patient. *Follow-up* and *Baseline* represent the image voxel values of the scans, *f(dose)* represent a function of the dose for that voxel and *LET* the linear energy transfer. Note that this model is the same as:

$$\text{Follow-up}_{i,j,n} = a \cdot \text{Baseline}_{i,n} + b \cdot \text{f}\left(\text{dose}_{i,n}\right) \cdot \left(1 + d \cdot \text{LET}_{i,n}\right) + \alpha_{i,j}$$

Where d=c/b. In other words, assuming LET=0 for photons gives RBE = 1 + c/b·LET. This is the form used for many pre-clinical studies, and one can therefore directly compare our results to published RBE models. To account for any potential non-linear effects of dose, we tested three different functional forms:

$$\begin{gathered} \mathrm{f}\left(\mathrm{dose}_{i,n}\right) = \mathrm{dose}_{i,n} \\ \mathrm{f}\left(\mathrm{dose}_{i,n}\right) = \ln\left(\mathrm{dose}_{i,n}\right) \\ \mathrm{f}\left(\mathrm{dose}_{i,n}\right) = {\mathrm{dose}_{i,n}}^{2} \end{gathered} \tag{2}$$

The $\alpha_{i,j}$ factor in equation (1) represents the random, mixed effects per patient and per follow-up session, accounting for heterogeneity of patients and scan sessions. Thus, the *b* and *c* indicate how much a change in image voxel value can be contributed to the dose or dose-LET interaction term, respectively; the model form corresponds a linear dependence of biological effect on LET. This model fully utilizes data from all patients and all time points in one fit.

The optimal model was subsequently determined by examining residuals for each model fit, but with preference to the linear model in close calls. Models for each modality - T1, T2 and FA - were fitted separately. The regression model with the optimal functional form for dose was used to predict image changes as a function of dose (spanning the dose range in 50 equal steps) with a fixed baseline voxel value (FA=0.26, T1=396 and T2=411, chosen as representative normal values for each scanning modality), five different fixed LETs (LET=0.9, 1.1, 1.5, 2.5 and 5 keV/µm) and 186 days to follow-up. To illustrate the impact of LET on imaging changes, we also compared predicted changes with varying LET values – but a constant dose – to the 'normal variation' in each of the imaging modalities. To obtain the range of 'normal variation', we created a mask for each patient of normal appearing white matter (NAWM) on the baseline scans. This mask was based on the FA signal, using the noise cut-off described in the data processing, but further excluded the GTV delineation. Mean and standard deviations of voxel values across all patients were acquired as estimates for NAWM for each modality.

To assess the inter- and intra-patient variation in dose and LET dependence over time, we conducted separate regressions for each follow-up session, i.e. using the adjusted model:

$$\text{Follow-up}_{i,j,n} = a_{i,j} \cdot \text{Baseline}_{i,n} + b_{i,j} \cdot \mathrm{f}\left(\mathrm{dose}_{i,n}\right) + c_{i,j} \cdot \mathrm{LET}_{i,n} \cdot \mathrm{f}\left(\mathrm{dose}_{i,n}\right) \tag{3}$$

Which provides a separate estimate for the *b* and *c* parameters for each patient and follow-up session. For simplicity, we only fitted this using the functional dependence of dose that was found optimal in the general model.

All data analysis was conducted in R (version 3.6.0) and all image registration and data preparation were done in Matlab (version 9.6.0.1135713 (R2019a) Update 3).

Results

A total of six patients were included between the ages of 6 and 17, who all received standard clinical proton therapy and four follow-up MRI sessions each. All patients were treated with passive scattering proton therapy. Further details on the patient treatment can be seen in Table 1.

Patient	Age	Gender	Cancer type	Previous treatment	Proton therapy treatment	Other treatment
1	12	F	Plexus choriodea carcinoma	Surgery + chemo	36 Gy 20 fractions	Adjuvant chemo
2	6	M	Ependymoma	Surgery + chemo	50.4 Gy + 3.6 Gy boost 28 fractions + 2 boost fractions	Chemo
3	9	F	Ependymoma	Surgery	59.4 Gy 33 fractions	None
4	15	M	Intracereberal geminoma	Stereotactic biopsy + chemo	24 Gy + 16 Gy boost 15 fractions + 10 boost fractions	None

5	15	M	Germcell tumor	Chemo + surgery + chemo	50.4 Gy + 3.6 Gy boost 28 fractions + 2 boost fractions	None
6	17	M	Bifocal intracerebral germinoma	Chemo	24 Gy + 16 Gy boost 15 fractions + 10 boost fractions	None

Table 1. An overview of the cancer type and treatment given to each patient. The prescription proton doses are given in Cobalt equivalent dose, which represents the physical proton dose multiplied with a relative biological effectiveness (RBE) of 1.1.

Patient 1 received MRI scans not adhering to the study protocol and was thus excluded from the analysis. Of the remaining five patients, there were issues with ten follow-up sessions. Seven had to be excluded due to missing data, and for three others the standard registration pathway failed and the T2 sequence had to be excluded or replaced from the analysis. An overview can be seen in Table 2. An example of our image registrations can be seen in Fig. 3.

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
FU 1 (~2 months after ended RT)	Ex	x	x	x	x	x*
FU 2 (~5 months after ended RT)	Ex	x(*)	x*	x	x	x
FU 3 (~8 months after ended RT)	Ex	x	x			
FU 4 (~11 months after ended RT)	Ex	x				

*Table 2. Available follow-up sessions for each patient. In each follow-up session a fractional anisotropy (FA) image, T1- and T2-weighted image is available. The * denotes sessions where the T2 image is not included in the deformable image registration, but for the session denoted with (*) another T2 that was acquired the same day on a PET/MRI scanner was successfully used as a replacement. "Ex" are scans that were excluded from the analysis because the MRI protocols were different from the dedicated study protocol MRIs. The blank spaces represent missing FA data.*

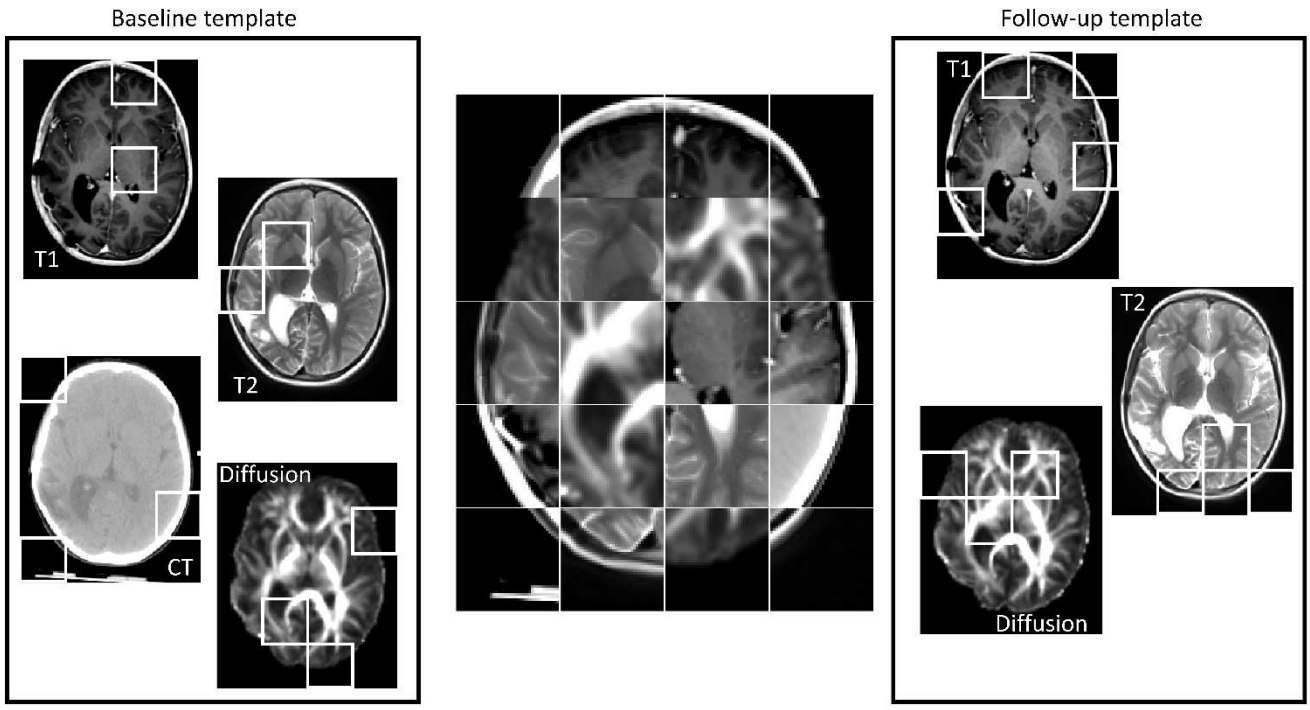


Figure 3. An example of the template deformable registration for one patient. In the middle is a mosaic plot of scans from the baseline template and a follow-up template. On the left, four baseline scans are shown as well as white squares to indicate which parts of the scans are shown in the middle mosaic. Similarly, on the right, scans from the follow-up template are shown with white squares corresponding to the locations in the mosaic plot

The linear function of dose (f(dose)=dose) was selected based on simplicity as minimal difference was observed in residual analysis. Fig. 3 show predictions of image changes as a function of dose and LET. This figure also shows our averaged NAWM values for each modality. We see a decrease in FA signal as a function of dose for all the investigated LET values. However, the signal increases as a function of LET, contrary to expectation[35–37], but all predictions are within the NAWM band. The T1 predictions appear relatively constant as a function of dose and LET, except for the very high-LET region. The T2 predictions show the largest dependence on both dose and LET, but with opposite effects; as the dose increases so does the T2 signal, but a higher LET value leads to a lower T2 signal.

In Table 3, the parameters *a*, *b* and *c* are shown for each modality. All three modalities show a relevant dose dependence through the *b* parameter as *b* is of the order of a change in voxel value that is impactful for each modality – FA per definition is always between 0 and 1. The *c* parameter is a factor 10 lower, compared to the *b* parameter, for FA and T1, while of the same order for the T2 model. This indicates the dose*LET-term is less impactful when predicting imaging changes as compared to the dose-term for FA and T1.

Prediction of image response as a function of dose for various LET values

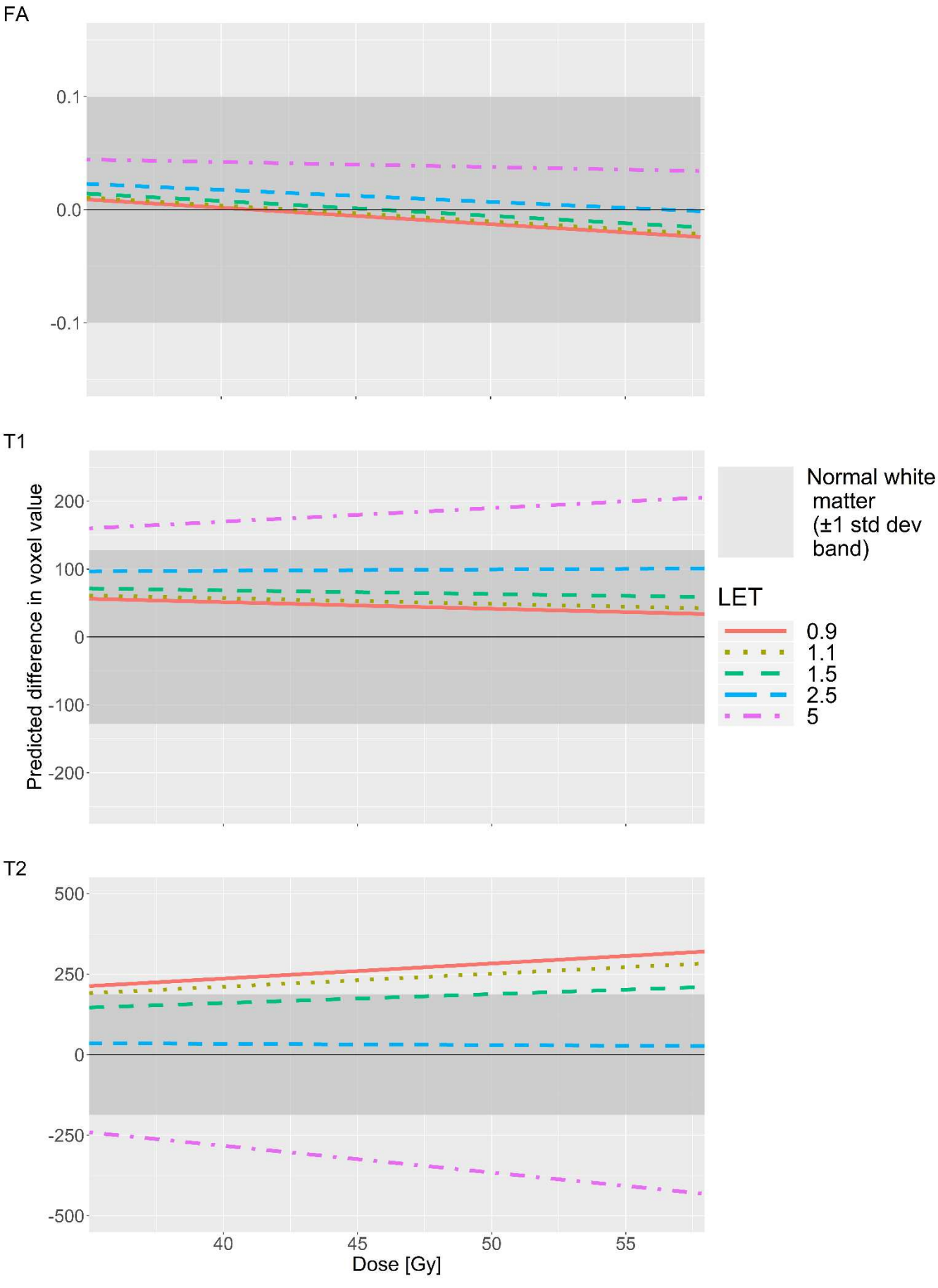


Figure 4. The predictions of model 1 (linear dose dependence), pooled over all patient and follow-up scan values as a function of dose for LET = [0.9, 1.1, 1.5, 2.5, 5.0] keV/ µm. The y-axis is defined such that a decrease here over time represents a lower voxel value in the follow-up scans as compared to baseline. This prediction assumes a representative voxel value of the baseline scan across all patients. The prediction is what is expected to happen for a theoretical patient at a follow-up after 186 days. The top figure shows the predicted values for fractional anisotropy (FA) MRI, the middle for T1-weighted MRI and the bottom figure is for T2-weighted MRI scan values.

Scan modality \ Model parameter	a (Baseline scan)	b (Dose)	c (Dose*LET)
FA	0.9973 ± 0.0006	$-1.37 \cdot 10^{-3} \pm 1 \cdot 10^{-5}$	$2.44 \cdot 10^{-4} \pm 5 \cdot 10^{-6}$
T1	0.8171 ± 0.0006	-1.47 ± 0.01	0.582 ± 0.004
T2	0.9042 ± 0.0008	7.23 ± 0.03	-2.45 ± 0.01

Table 3. Parameter a, b and c after modelling Eq (1) using the complete nested dataset for a linear dependence on dose. The parameters are given separately for each of the three modalities considered, and with one standard deviation.

We additionally predicted voxel value change for various LET values and a fixed dose to the voxel of 35 Gy for FA and T2, as seen in Fig. 4. For the T2 image, a continuous decrease in signal was seen and this signal was different from NAWM for LET > ~ 5.5 keV/µm. We investigated the predictions shown in Table 3, Fig. 4 and Fig. 5 with different dose thresholds (5 and 20 Gy) with slightly varying results, see the summary in the supplementary material and Supplementary Figure 2-5 and Supplementary Table 1-2.

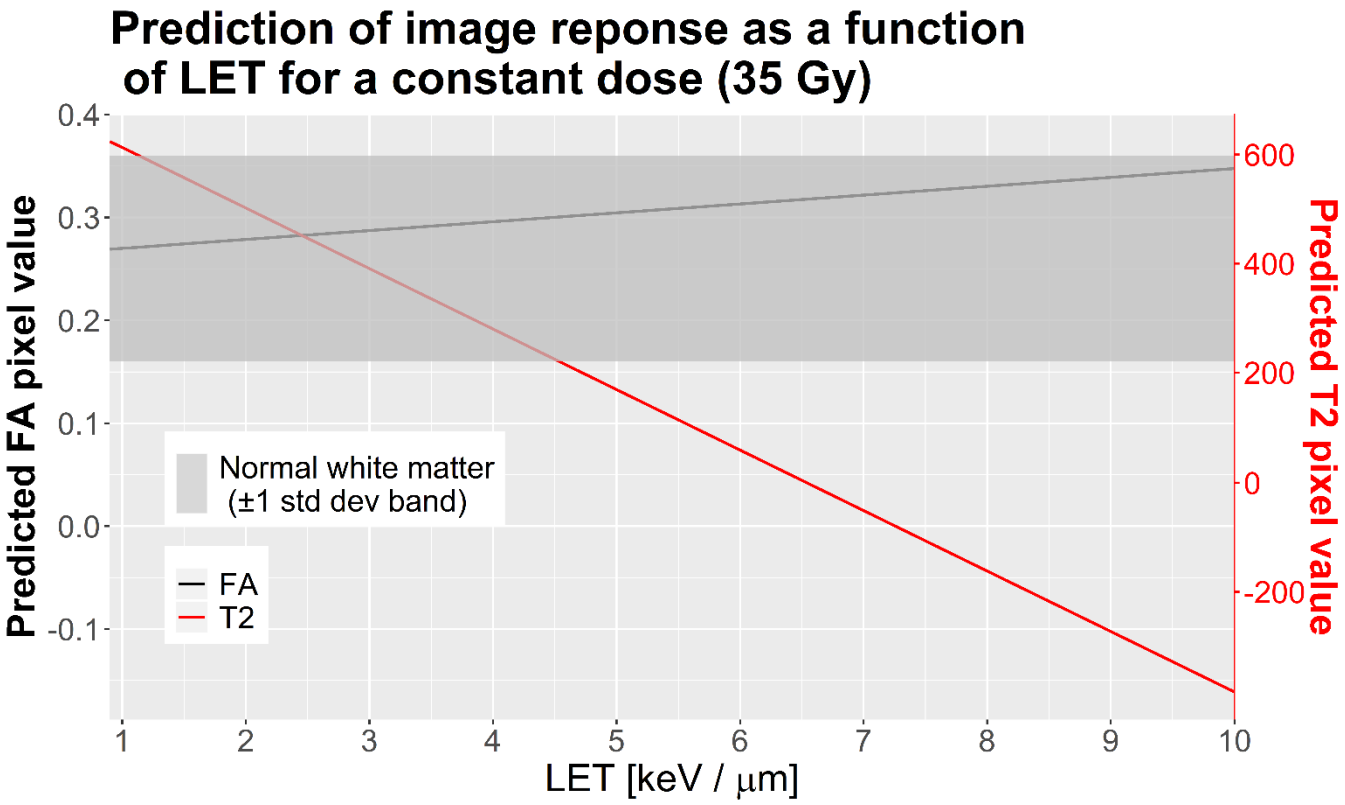


Figure 5. The predicted image change based on a linear dependence of dose for modalities fractional anisotropy (FA) (y-axis on the left) and T2-weighted MRI (y-axis on the right) as a function of linear energy transfer (LET). The predictions are made at 35 Gray and for a standard voxel value at baseline (FA=0.26, T2=411). The grey band indicate average normal appearing white matter value for both modalities with 1 standard deviation.

The fit parameters *b* and *c* from equation (3) for individual follow-up sessions can be seen in Fig. 5. Here, points below the zero line indicate a decrease in voxel value over time, and points above indicate an increase over time. The y-axis is scaled to facilitate comparison between *b* and *c*: The limits on the *y*-axis are found as the minimal and maximal value of *b* and *c* after they have been multiplied by the (median dose) and (median dose*median LET) respectively. This way, the y-scale becomes a reflection of the median change in FU value compared to baseline that arises due to the (dose)- and (dose*LET)-term in the voxels. This normalization is done internally for each modality. Intra patient variations seemed small compared to inter-patient variations for all modalities as seen by the grouping internally for each patient among the same parameter value. The b-values are negative for the FA image indicating the expected

decrease in FA as function of dose. The *c*-values appear inconclusive and mostly smaller compared to the variations of the *b*-values, indicating that this dose*LET term did not have as great an impact on imaging changes as the dose. No clear conclusions or consistency could be seen in the T1 image, which was a general trend across all analysis made in this study.

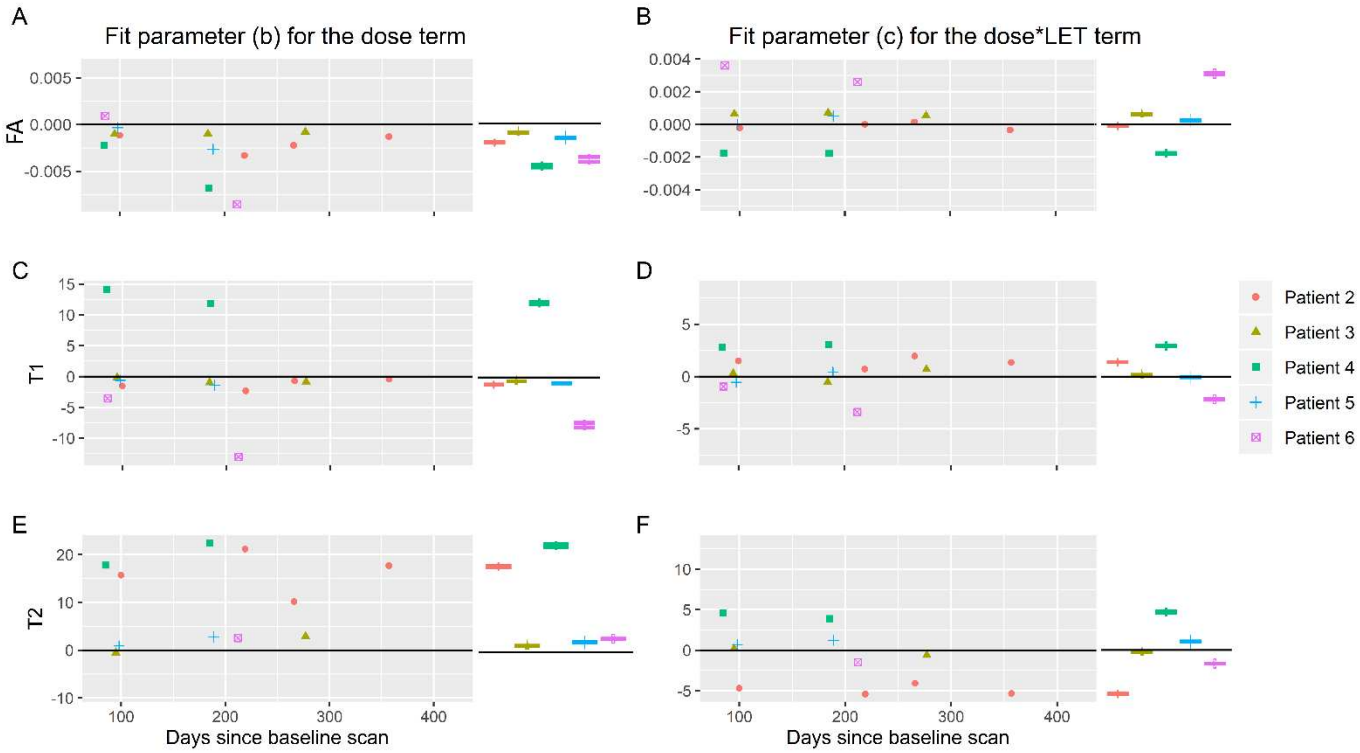


Figure 6. Fit parameters b *and* c *to model (1), modelling the effect of dose and dose*LET, respectively (LET=linear energy transfer). These parameters are shown for all patients and all follow-ups as a function of days since baseline scan and for fractional anisotropy (FA) (top row), T1-weighted MRI (middle row) and T2-weighted MRI (bottom row) scans. To the right of each subplot, the* b *and* c *parameters are shown with error bars (one standard deviation) from five individual patient models; one for each patient including all follow-up data from that patient. The range of the y-axis has been adjusted to an equal range between all plots in terms of impact on the follow-up scan value in equation (3).*

Discussion

3D imaging assays of radiation damage are promising to derive human tissue specific biological effects of radiation as function of dose and LET. The 3D in-vivo assay allows sampling across multiple dose levels and LET values in the same patient, thus avoiding the need to accrue large numbers of patients to observe the signal from dose and LET over the random variation across patients. Continuous predictors typically also have much more statistical power than a binary time-to-event endpoint of clinical manifest toxicity.

On the downside, longitudinal imaging studies are hard to analyze in terms of image registration and in terms of statistical analysis of the high dimensionality data. Here we have tried to tackle those two problems by providing a framework for using longitudinal deformable image registrations and voxel-based statistical analysis, accounting for nesting of data on patient and follow-up session level, to assess the effect of dose and LET on the image assays across several patients and time points in each patient. Unfortunately, the number of patients accrued in our project is too small to draw firm conclusions. However, the expected dose response is derived on FA maps and we do provide means to assess the importance of LET compared to dose and interpatient variation. We observe substantial patient to patient variation in the internal model fits (i.e. sensitivity to dose and LET); in other words, inter-patient variation in radiation sensitivity may be a stronger factor than variation in biological effect across clinical proton beams.

A handful of previous publications have examined the use of imaging changes as markers for radiation damage to assess biological effectiveness of proton beam irradiation as function of dose and LET[14,38–40]. In particular, two recent papers reported on radiation-induced brain lesions in adult patients, and related the presence of imaging changes to dose and LET on a voxel-by-voxel basis[39,40]. The paper by Bahn et al demonstrated a relationship between imaging changes and LET in 110 patients treated for low-grade glioma, while Niemierko et al were unable to find a LET dependence of necrotic changes in fifty patients treated for head and neck, skull base, or intracranial tumors. Conversely, they concluded that inter-patient variability in radio sensitivity was likely to overshadow any LET effect; much in line with our present

findings. Of note, neither of the two studies used deformable image registration or assessed imaging changes as continuous measures.

We have deliberately chosen a formal statistical methodology testing the effect of dose and LET*dose using conventional statistical methods rather than some of the widely available models of RBE rooted in mechanistic or in-vitro modeling. The use of the pure statistical method here should ideally be compared to mechanistic modeling in larger series, but in terms of advantages the current method avoids the need to combine LQ modeling with RBE using formulas which may be stretched beyond clinical utility when applied to human tissue. Note, however, that our relatively simple functional form is equivalent to the linear dependence of RBE on LET used for a number of models based on pre-clinical data: RBE = 1+K*LET (where K=c/b for the current model). The ratio between the c and b parameters reported here thus provides an estimate of the dependence of the relative biological effectiveness on LET. Comparison using the methods proposed here could hopefully shed light on the question of the modifying effect of LET on the dose in proton therapy in future series. Other weaknesses in our data analysis include not taking fraction size effects into account, averaging dose and LET over multi-phase treatments for some patients, and assuming independence of the radiation response for adjacent voxels.

In this study we rely heavily on imaging data and changes to measure a response, but many of the patients included in the study did not exhibit significant image changes during our follow-up period. This could explain the diverging results. Also, the correlation between image changes and clinically meaningful endpoints is often weak and the functional form not well established. For example, our choice of FA signal as imaging marker has been questioned since protocol initiation: A study from 2019[41] showed an increase in FA signal due to radiation damage and it was hypothesized that this increase was driven by an overall reduction in mean diffusivity. Another study[23] showed that different white matter tracts are differently susceptible to radiation damage, with some tracts remaining unchanged, some showing an increase in FA signal, and some showing a decrease in FA signal. Additionally, the T1/T2 masks used in this study are very

general and include both white matter, grey matter and organs at risk, all of which can have very different responses to radiation that cannot be captured by our model. This points to a potential critical weakness in the imaging assays: If the imaging marker is a composite signal of unspecific changes due to radiation, it may fail to provide guidance to the relative effect of radiation as function of LET. However, given a sufficiently strong imaging marker, our framework can enable in-vivo studies of LET effects from radiation for specific tissues and/or regions. One could investigate specific structures, such as the brain ventricles, and create radiobiological models that account for intra-patient variations through our proposed regression analysis. We strongly suggest selecting imaging biomarkers which has been associated with a relevant clinical outcome. This framework should also be extended to include neighboring-voxel effects[42] to improve accuracy of the models.

The deformable registration part of the framework can be applied to almost any follow-up analysis without the need to fine-tune the algorithm. It uses standard values for the deformable registrations and is thus generalizable to other datasets from other centers. Although there are no requirements on the resolution of the included scans, to compare the results across multi-center datasets, we recommend using standardized protocols to avoid inconsistencies and inhomogeneities.

Conclusions

We have demonstrated a framework for voxel-based registration and phenomenological fitting of longitudinal image data versus dose and LET to derive estimates of the biological effect of high LET radiation. This framework is available online[1].

Acknowledgements

This work has been supported by research grants from Børnecancerfonden (The Danish Childhood Cancer Foundation), grant numbers 2013-33, 2014-34 and 2015-47. Ane Appelt is supported by Yorkshire Cancer Research Academic Fellowship funding (grant L389AA).

Conflict of interest

The authors have no conflicts of interests to disclose.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. Skaarup M. Longitudinal image analysis of biological effects. 2021. doi:10.5281/zenodo.4569646

2. Hall N, Rouge B. The physics of proton therapy. *Phys Med Biol*. 2015;60(8). doi:10.1088/0031-9155/60/8/R155.The

3. Mohan R, Grosshans D. Proton therapy – Present and future. *Adv Drug Deliv Rev*. 2017;109:26-44. doi:10.1016/j.addr.2016.11.006

4. Holloway SM, Holloway MD, Thomas SJ. A method for acquiring random range uncertainty probability distributions in proton therapy. *Phys Med Biol*. 2018;63(1):aa9502. doi:10.1088/1361-6560/aa9502

5. Taasti VT, Jeong J, Jackson A, Deasy JO. A theoretical investigation of adequate range uncertainty margins in proton treatment planning to preserve tumor control probability. *Acta Oncol (Madr)*. 2019;58(10):1446-1450. doi:10.1080/0284186X.2019.1627415

6. Malyapa R, Lowe M, Bolsi A, Lomax AJ, Weber DC, Albertini F. Evaluation of Robustness to Setup and Range Uncertainties for Head and Neck Patients Treated with Pencil Beam Scanning Proton Therapy. *Int J Radiat Oncol Biol Phys*. 2016;95(1):154-162. doi:10.1016/j.ijrobp.2016.02.016

7. Liebl J, Paganetti H, Zhu M, Winey BA. The influence of patient positioning uncertainties in proton radiotherapy on proton range and dose distributions. *Med Phys*. 2014;41(9):1-12.

doi:10.1118/1.4892601

8. Paganetti H. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Phys Med Biol*. 2014;59(22):R419-R472. doi:10.1088/0031-9155/59/22/R419

9. McNamara AL, Schuemann J, Paganetti H. A phenomenological relative biological effectiveness (RBE) model for proton therapy based on all published in vitro cell survival data. *Phys Med Biol*. 2015;60(21):8399-8416. doi:10.1088/0031-9155/60/21/8399

10. Wilkens JJ, Oelfke U. A phenomenological model for the relative biological effectiveness in therapeutic proton beams. *Phys Med Biol*. 2004;49(13):2811-2825. doi:10.1088/0031-9155/49/13/004

11. McMahon SJ, Paganetti H, Prise KM. LET-weighted doses effectively reduce biological variability in proton radiotherapy planning. *Phys Med Biol*. 2018. doi:10.1088/1361-6560/aae8a5

12. Dahle TJ, Rykkelid AM, Stokkevåg CH, et al. Monte Carlo simulations of a low energy proton beamline for radiobiological experiments. *Acta Oncol (Madr)*. 2017. doi:10.1080/0284186X.2017.1289239

13. Underwood TSA, Grassberger C, Bass R, et al. Asymptomatic Late-phase Radiographic Changes Among Chest-Wall Patients Are Associated With a Proton RBE Exceeding 1.1. *Int J Radiat Oncol Biol Phys*. 2018. doi:10.1016/j.ijrobp.2018.03.037

14. Peeler CR, Mirkovic D, Titt U, et al. Clinical evidence of variable proton biological effectiveness in pediatric patients treated for ependymoma. *Radiother Oncol*. 2016. doi:10.1016/j.radonc.2016.11.001

15. Wang CC, McNamara AL, Shin J, et al. End-of-Range Radiobiological Effect on Rib Fractures in

Patients Receiving Proton Therapy for Breast Cancer. *Int J Radiat Oncol Biol Phys*. 2020;107(3):449-454. doi:10.1016/j.ijrobp.2020.03.012

16. Chen Y, Ahmad S. Empirical model estimation of relative biological effectiveness for proton beam therapy. *Radiat Prot Dosimetry*. 2012;149(2):116-123. doi:10.1093/rpd/ncr218

17. Jones KM, Michel KA, Bankson JA, Fuller CD, Klopp AH, Venkatesan AM. Emerging Magnetic Resonance Imaging Technologies for Radiation Therapy Planning and Response Assessment. *Int J Radiat Oncol Biol Phys*. 2018;101(5):1046-1056. doi:10.1016/j.ijrobp.2018.03.028

18. Robbins ME, Brunso-Bechtold JK, Peiffer AM, Tsien CI, Bailey JE, Marks LB. Imaging Radiation-Induced Normal Tissue Injury. *Radiat Res*. 2012;177(4):449-466. doi:10.1667/rr2530.1

19. Minkoff D, Gill BS, Kang J, Beriwal S. Cervical cancer outcome prediction to high-dose rate brachytherapy using quantitative magnetic resonance imaging analysis of tumor response to external beam radiotherapy. *Radiother Oncol*. 2015;115(1):78-83. doi:10.1016/j.radonc.2015.03.007

20. Bhatnagar P, Subesinghe M, Patel C, Prestwich R, Scarsbrook AF. Functional imaging for radiation treatment planning, response assessment, and adaptive therapy in head and neck cancer. *Radiographics*. 2013;33(7):1909-1929. doi:10.1148/rg.337125163

21. McVicar N, Popescu IA, Heath E. Techniques for adaptive prostate radiotherapy. *Phys Medica*. 2016;32(3):492-498. doi:10.1016/j.ejmp.2016.03.010

22. Connor M, Karunamuni R, McDonald C, et al. Dose-dependent white matter damage after brain radiotherapy. *Radiother Oncol*. 2016;121(2):209-216. doi:10.1016/j.radonc.2016.10.003

23. Connor M, Karunamuni R, McDonald C, et al. Regional susceptibility to dose-dependent white matter damage after brain radiotherapy. *Radiother Oncol*. 2017;123(2):209-217. doi:10.1016/j.radonc.2017.04.006

24. Haris M, Kumar S, Raj MK, et al. Serial diffusion tensor imaging to characterize radiation-induced changes in normal-appearing white matter following radiotherapy in patients with adult low-grade gliomas. *Radiat Med - Med Imaging Radiat Oncol*. 2008;26(3):140-150. doi:10.1007/s11604-007-0209-4

25. Palmer SL, Glass JO, Li Y, et al. White matter integrity is associated with cognitive processing in patients treated for a posterior fossa brain tumor. *Neuro Oncol*. 2012;14(9):1185-1193. doi:10.1093/neuonc/nos154

26. Uh J, Merchant TE, Conklin HM, et al. Diffusion tensor imaging-based analysis of baseline neurocognitive function and post-treatment white matter changes in pediatric patients with craniopharyngioma treated with surgery and proton therapy. *Int J Radiat Oncol*. 2020;S0360(3016):34219-X. doi:10.1016/j.colsurfa.2020.124530

27. University Hospital Muenster. Prospective Trial for the Diagnosis and Treatment of Intracranial Germ Cell Tumors (SIOPCNSGCTII). clinicaltrials.gov [cited 2021 Feb 1]. Available from: https://clinicaltrials.gov/ct2/show/record/NCT01424839.

28. Centre Leon Berard. An International Clinical Program for the Diagnosis and Treatment of Children With Ependymoma (SIOP-EP-II) - Protocol version 2.1 from 2015 May 29th. clinicaltrials.gov [cited 2021 Feb 1]. Available from: https://clinicaltrials.gov/ct2/show/NCT02265770.

29. Marner L, Nysom K, Sehested A, et al. Early postoperative 18F-FET PET/MRI for pediatric brain and spinal cord tumors. *J Nucl Med*. 2019;60(8):1053-1058. doi:10.2967/jnumed.118.220293

30. Titt U, Sahoo N, Ding X, et al. Assessment of the accuracy of an MCNPX-based Monte Carlo simulation model for predicting three-dimensional absorbed dose distributions. *Phys Med Biol*. 2008;53(16):4455-4470. doi:http://dx.doi.org/10.1088/0031-9155/53/16/016

31. Sawakuchi GO, Mirkovic D, Perles LA, et al. An MCNPX Monte Carlo model of a discrete spot scanning proton beam therapy nozzle. *Med Phys*. 2010;37(9):4960-4970. doi:10.1118/1.3476458

32. Darkner S, Sporring J. Locally orderless registration. *IEEE Trans Pattern Anal Mach Intell*. 2013. doi:10.1109/TPAMI.2012.238

33. Darkner S, Pai A, Liptrot MG, Sporring J. Collocation for Diffeomorphic Deformations in Medical Image Registration. *IEEE Trans Pattern Anal Mach Intell*. 2018. doi:10.1109/TPAMI.2017.2730205

34. Omar RZ, Wright EM, Turner RM, Thompson SG. Analysing repeated measurements data: A practical comparison of methods. *Stat Med*. 1999;18(13):1587-1603. doi:10.1002/(SICI)1097-0258(19990715)18:13<1587::AID-SIM141>3.0.CO;2-Z

35. Khong PL, Leung LHT, Fung ASM, et al. White matter anisotropy in post-treatment childhood cancer survivors: Preliminary evidence of association with neurocognitive function. *J Clin Oncol*. 2006;24(6):884-890. doi:10.1200/JCO.2005.02.4505

36. Chapman CH, Nazem-Zadeh M, Lee OE, et al. Regional Variation in Brain White Matter Diffusion Index Changes following Chemoradiotherapy: A Prospective Study Using Tract-Based Spatial Statistics. *PLoS One*. 2013;8(3). doi:10.1371/journal.pone.0057768

37. Uh J, Merchant TE, Li Y, et al. Differences in brainstem fiber tract response to radiation: A longitudinal diffusion tensor imaging study. *Int J Radiat Oncol Biol Phys*. 2013;86(2):292-297. doi:10.1016/j.ijrobp.2013.01.028

38. Eulitz J, Troost EGC, Raschke F, et al. Predicting late magnetic resonance image changes in glioma patients after proton therapy. *Acta Oncol (Madr)*. 2019;58(10):1536-1539. doi:10.1080/0284186X.2019.1631477

39. Bahn E, Bauer J, Harrabi S, Herfarth K, Debus J, Alber M. Late Contrast Enhancing Brain Lesions in

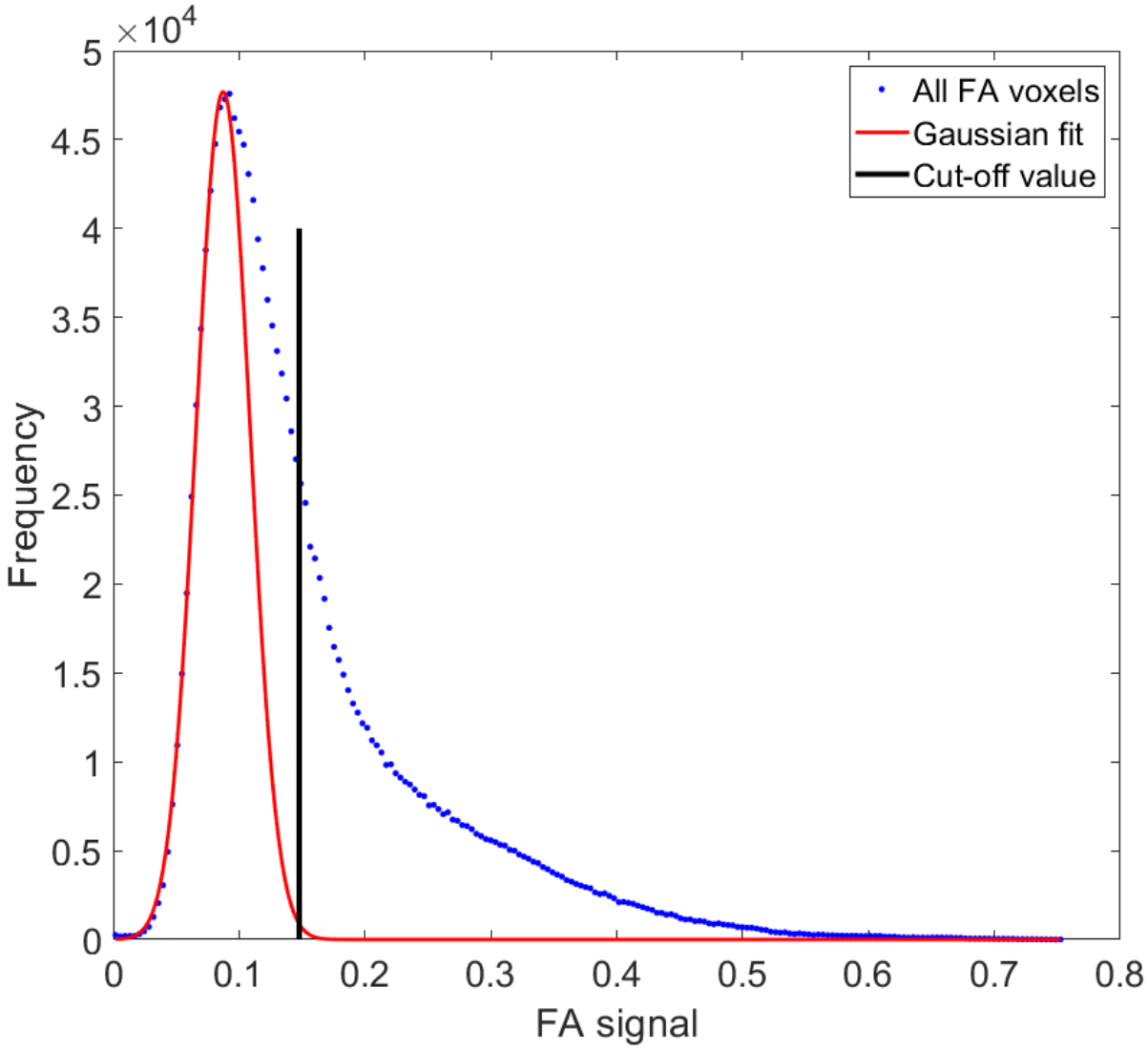
Proton-Treated Patients With Low-Grade Glioma: Clinical Evidence for Increased Periventricular Sensitivity and Variable RBE. *Int J Radiat Oncol Biol Phys*. 2020;107(3):571-578. doi:10.1016/j.ijrobp.2020.03.013

40. Andrzej Niemierko, Schuemann J, Niyazi M, et al. Brain necrosis in adult patients after proton therapy: Is there evidence for dependency on linear energy transfer (LET)? *Int J Radiat Oncol Biol Phys*. 2020:141926. doi:http://dx.doi.org/10.1016/j.ijrobp.2020.08.058

41. Raschke F, Wesemann T, Wahl H, et al. Reduced diffusion in normal appearing white matter of glioma patients following radio(chemo)therapy. *Radiother Oncol*. 2019;140:110-115. doi:10.1016/j.radonc.2019.06.022

42. Bowen SR, Hippe DS, Art Chaovalitwongse W, et al. Voxel forecast for precision oncology: Predicting spatially variant and multiscale cancer therapy response on longitudinal quantitative molecular imaging. *Clin Cancer Res*. 2019;25(16):5027-5037. doi:10.1158/1078-0432.CCR-18-3908

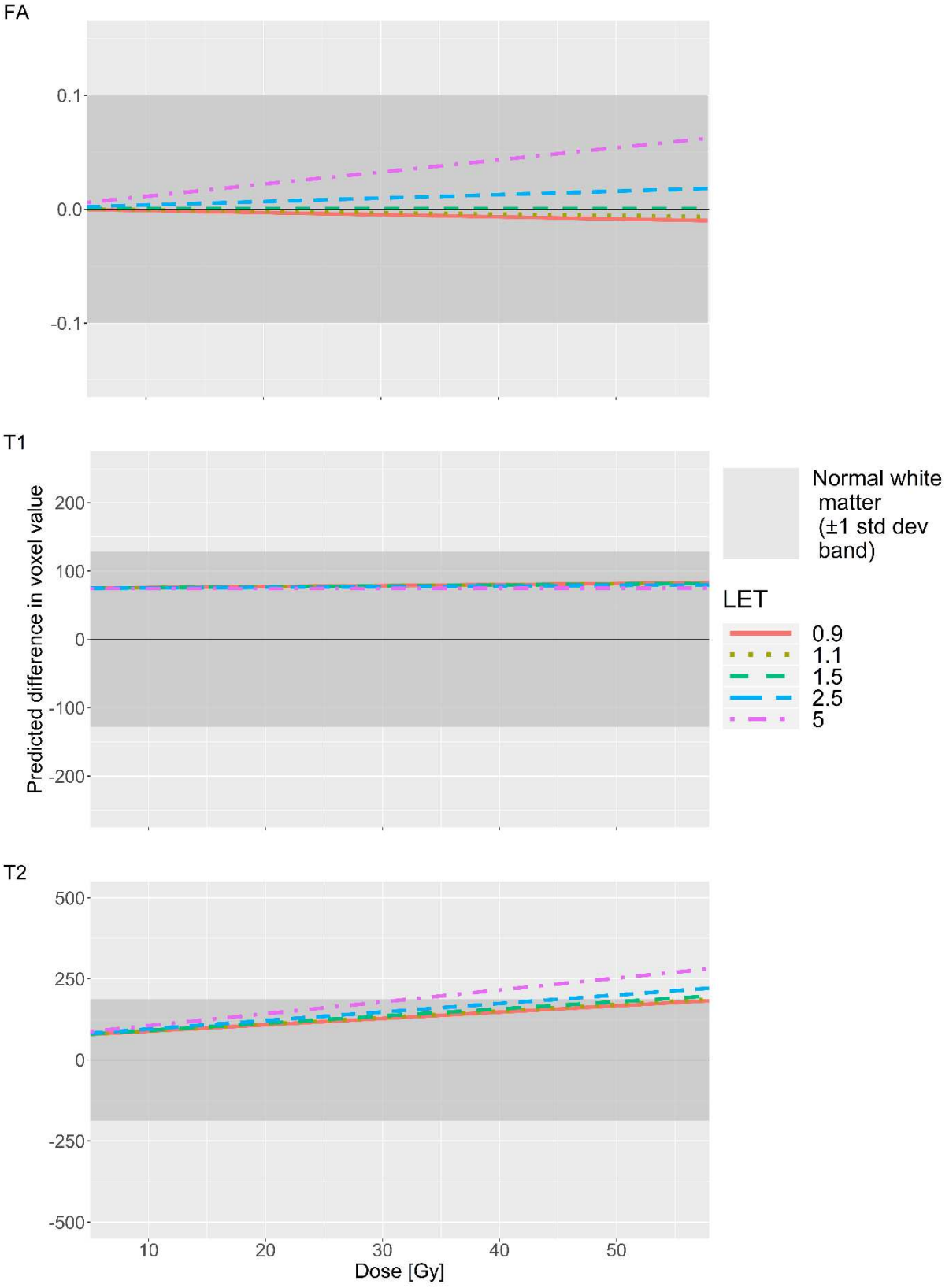
Supplementary Material



Supplementary Figure 1. A histogram of the baseline fractional anisotropy (FA) voxel values for patient 2. The red line is a gaussian fit to the bins from 0 to the maximal frequency bin of the histogram. It is used to find the noise level of the FA scan. The black line is the cut-off value used, defined by the mean of the gaussian fit plus two standard deviations.

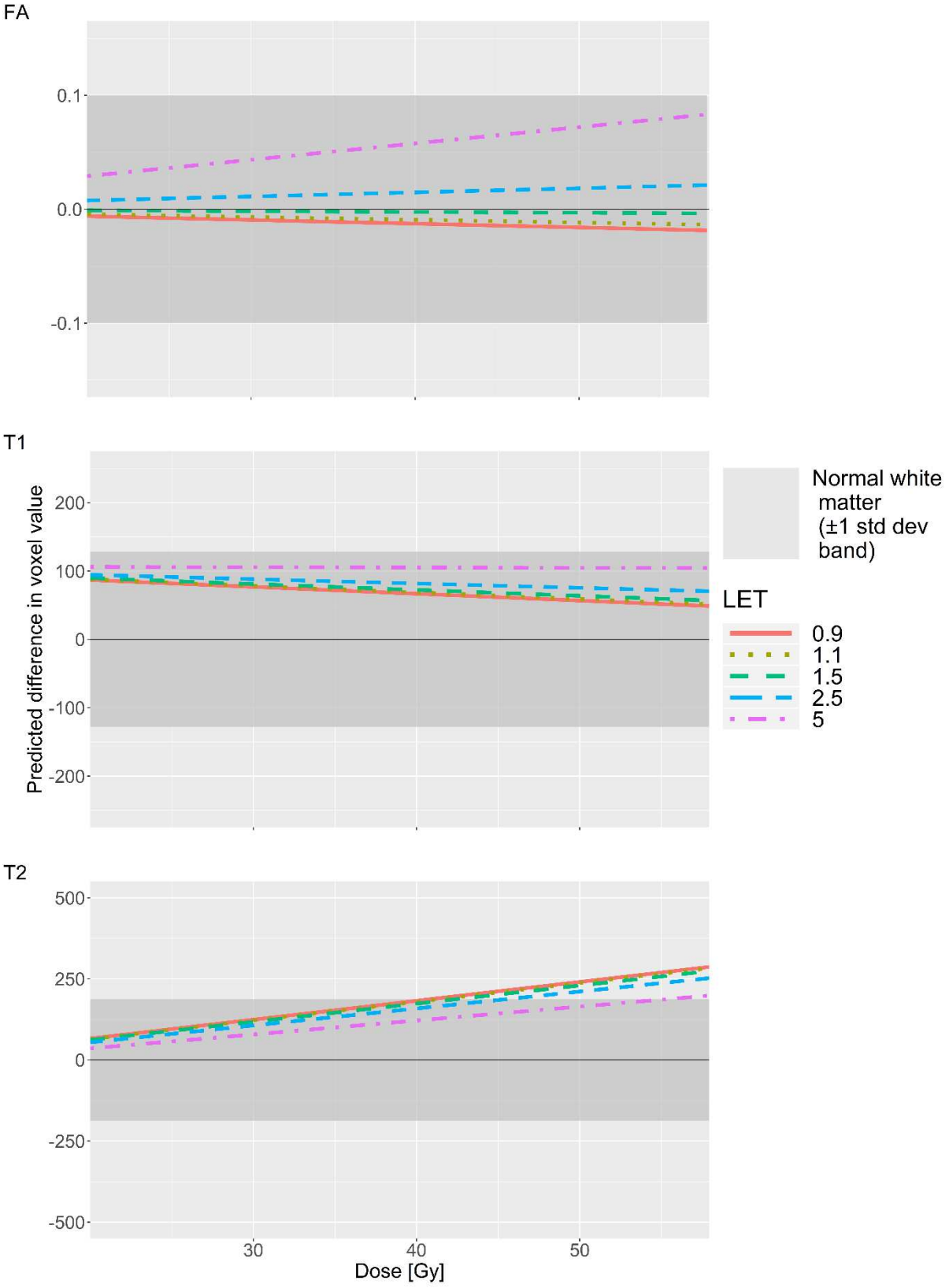
To investigate how the dose thresholds affected the results presented, we repeated the analysis including all voxels receiving more than 5 and 20 Gy into the model, respectively. For the 5 Gy threshold, all imaging changes were within the NAWM band and thus the model predicted no imaging changes for voxels receiving >5 Gy (Supplementary Figure 2 and 4). Similarly, parameter *b* and *c* of the model were lower across all modalities, indicating a small impact on imaging changes from the dose- and dose*LET -term (Supplementary Table 1). When we included all voxel receiving 20 Gy or more into the model, the predicted imaging changes were again within the NAWM-band (Supplementary Figure 3 and 5). Parameter *b* for a 20 Gy dose cut-off was slightly lower for all modalities (Supplementary Table 2) as compared to the values presented in Table 3, while parameter *c* seemed to have a slightly larger impact on imaging changes for FA and T2 for the 20 Gy cut-off.

Prediction of image response as a function of dose for various LET values

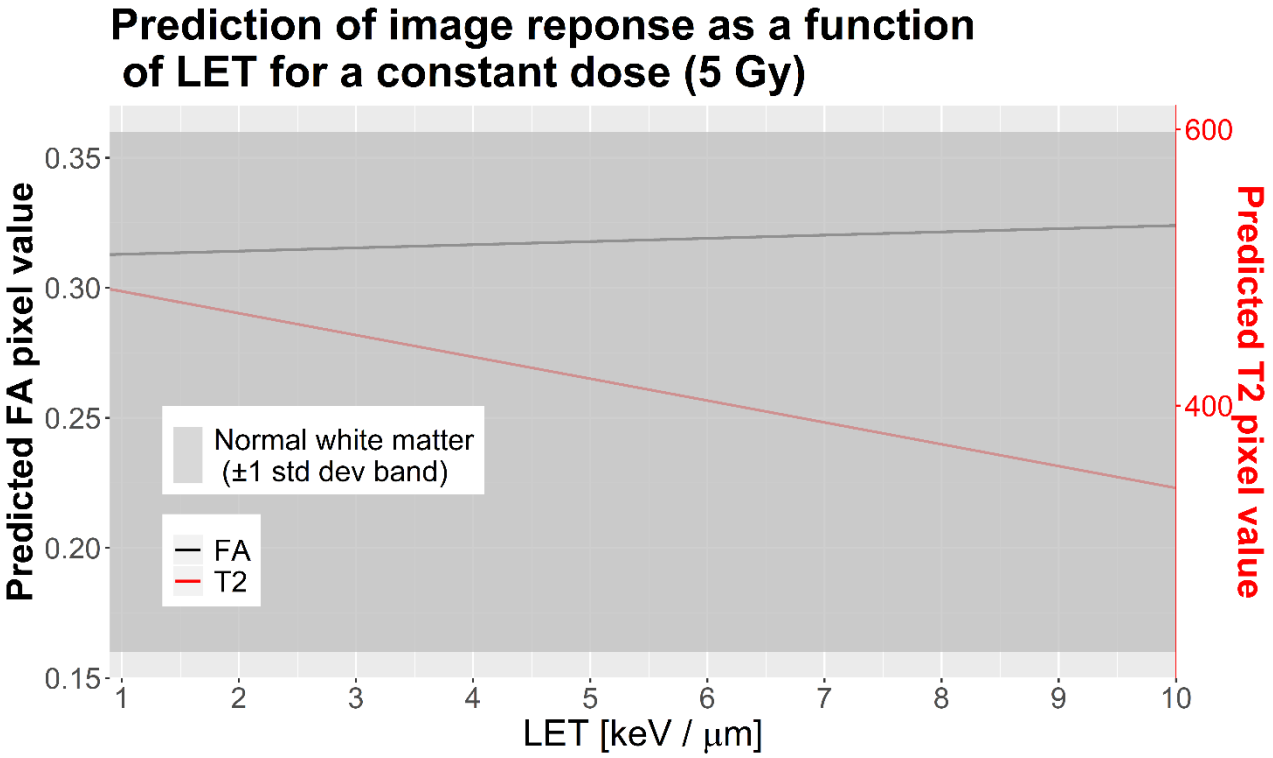


Supplementary Figure 2. The predictions of model 1 (linear dose dependence), pooled over all patient and follow-up scan values as a function of dose for LET = [0.9, 1.1, 1.5, 2.5, 5.0] keV/ µm. All voxels receiving 5 Gy or more were included in the model. The y-axis is defined such that a decrease here over time represents a lower voxel value in the follow-up scans as compared to baseline. This prediction assumes a representative voxel value of the baseline scan across all patients. The prediction is what is expected to happen for a theoretical patient at a follow-up after 186 days. The top figure shows the predicted values for fractional anisotropy (FA) MRI, the middle for T1-weighted MRI and the bottom figure is for T2-weighted MRI scan values.

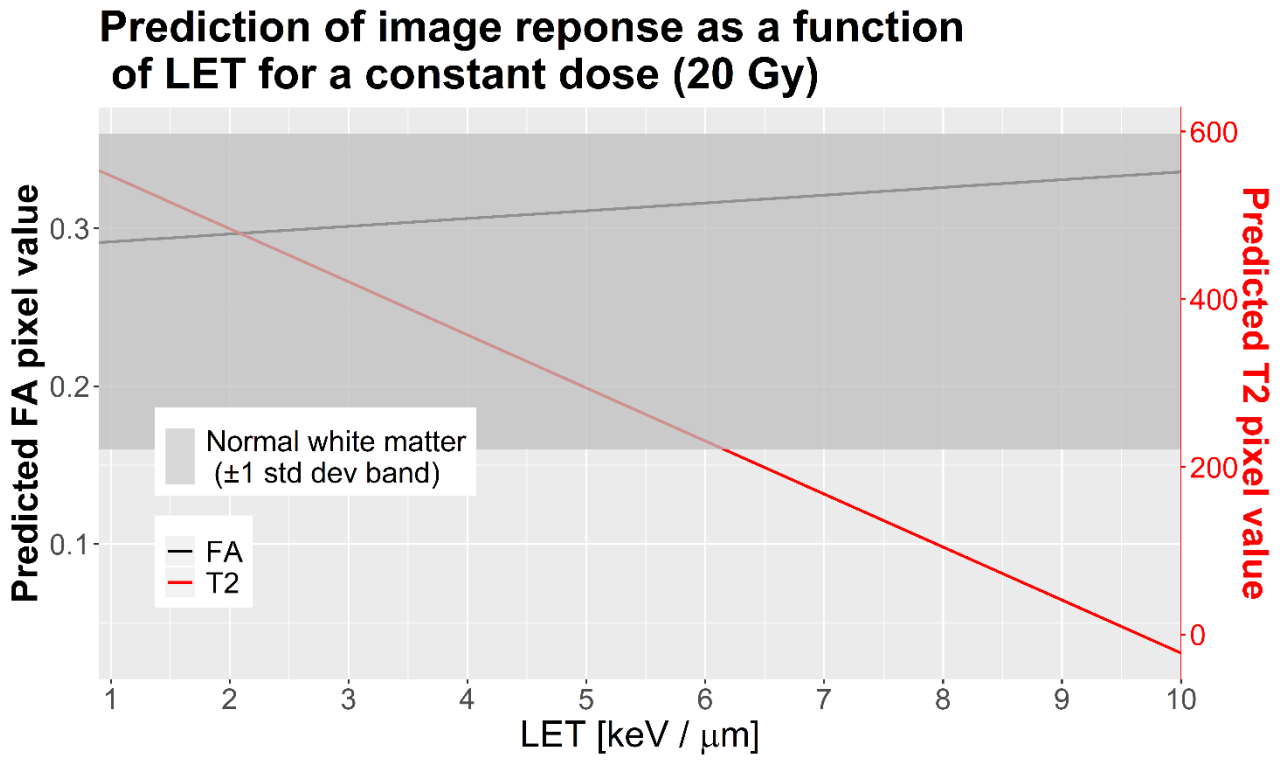
Prediction of image response as a function of dose for various LET values



Supplementary Figure 3. The predictions of model 1 (linear dose dependence), pooled over all patient and follow-up scan values as a function of dose for LET = [0.9, 1.1, 1.5, 2.5, 5.0] keV/ µm. All voxels receiving 20 Gy or more were included in the model. The y-axis is defined such that a decrease here over time represents a lower voxel value in the follow-up scans as compared to baseline. This prediction assumes a representative voxel value of the baseline scan across all patients. The prediction is what is expected to happen for a theoretical patient at a follow-up after 186 days. The top figure shows the predicted values for fractional anisotropy (FA) MRI, the middle for T1-weighted MRI and the bottom figure is for T2-weighted MRI scan values.



Supplementary Figure 4. The predicted image change based on a linear dependence of dose for modalities fractional anisotropy (FA) (black, y-axis on the left) and T2-weighted MRI (red, y-axis on the right) as a function of linear energy transfer (LET). All voxels receiving 5 Gy or more were included in the model. The predictions are made at 5 Gy and for a standard voxel value at baseline (FA=0.26, T2=411). The grey band indicate average normal appearing white matter value for both modalities with 1 standard deviation.



Supplementary Figure 5. The predicted image change based on a linear dependence of dose for modalities fractional anisotropy (FA) (black, y-axis on the left) and T2-weighted MRI (red, y-axis on the right) as a function of linear energy transfer (LET). All voxels receiving 20 Gy or more were included in the model. The predictions are made at 20 Gy and for a standard voxel value at baseline (FA=0.26, T2=411). The grey band indicate average normal appearing white matter value for both modalities with 1 standard deviation.

Scan modality \ Model parameter	a (Baseline scan)	b (Dose)	c (Dose*LET)
FA	0.9873 ± 0.0002	$-4.61 \cdot 10^{-4} \pm 4 \cdot 10^{-6}$	$3.05 \cdot 10^{-4} \pm 2 \cdot 10^{-6}$
T1	0.8451 ± 0.0003	0.178 ± 0.005	-0.033 ± 0.002
T2	-0.8054 ± 0.0002	1.58 ± 0.01	0.416 ± 0.004

Supplementary Table 1. Parameter a, b and c after modelling Eq (1) using the complete nested dataset for a linear dependence on dose. The data in the model includes all voxels receiving a dose above 5 Gy. The parameters are given separately for each of the three modalities considered, and with one standard deviation.

Supplementary Table 2

Scan modality \ Model parameter	a (Baseline scan)	b (Dose)	c (Dose*LET)
FA	0.9913 ± 0.0003	$-7.15 \cdot 10^{-4} \pm 7 \cdot 10^{-6}$	$4.30 \cdot 10^{-4} \pm 3 \cdot 10^{-6}$
T1	0.9006 ± 0.0004	-1.22 ± 0.01	0.234 ± 0.003
T2	0.7926 ± 0.0004	6.15 ± 0.02	-3.72 ± 0.01

Supplementary Table 2. Parameter a, b and c after modelling Eq (1) using the complete nested dataset for a linear dependence on dose. The data in the model includes all voxels receiving a dose above 20 Gy. The parameters are given separately for each of the three modalities considered, and with one standard deviation.