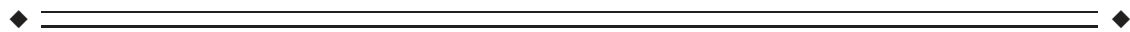


Four In Vivo *g*-Ratio-Weighted Imaging Methods: Comparability and Repeatability at the Group Level

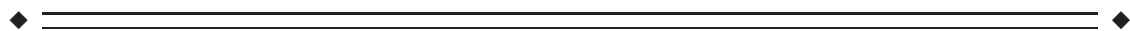
Isabel Ellerbrock * and Siawoosh Mohammadi

Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf,
Hamburg, Germany



Abstract: A recent method, denoted in vivo *g*-ratio-weighted imaging, has related the microscopic *g*-ratio, only accessible by ex vivo histology, to noninvasive MRI markers for the fiber volume fraction (FVF) and myelin volume fraction (MVF). Different MRI markers have been proposed for *g*-ratio weighted imaging, leaving open the question which combination of imaging markers is optimal. To address this question, the repeatability and comparability of four *g*-ratio methods based on different combinations of, respectively, two imaging markers for FVF (tract-fiber density, TFD, and neurite orientation dispersion and density imaging, NODDI) and two imaging markers for MVF (magnetization transfer saturation rate, MT, and, from proton density maps, macromolecular tissue volume, MTV) were tested in a scan–rescan experiment in two groups. Moreover, it was tested how the repeatability and comparability were affected by two key processing steps, namely the masking of unreliable voxels (e.g., due to partial volume effects) at the group level and the calibration value used to link MRI markers to MVF (and FVF). Our data showed that repeatability and comparability depend largely on the marker for the FVF (NODDI outperformed TFD), and that they were improved by masking. Overall, the *g*-ratio method based on NODDI and MT showed the highest repeatability (90%) and lowest variability between groups (3.5%). Finally, our results indicate that the calibration procedure is crucial, for example, calibration to a lower *g*-ratio value ($g = 0.6$) than the commonly used one ($g = 0.7$) can change not only repeatability and comparability but also the reported dependency on the FVF imaging marker. *Hum Brain Mapp* 39:24–41, 2018. © 2017 Wiley Periodicals, Inc.

Key words: *g*-ratio; repeatability; comparability; multiparameter mapping; diffusion-weighted imaging



INTRODUCTION

The myelin sheaths of an axon act as an electrical insulator. The *g*-ratio, the ratio of the inner to the outer diameter of the myelin sheath of a myelinated axon (or fiber), quan-

tifies the relative myelination of an axon and is, thus, an indicator of its conduction velocity [Waxman, 1980]. A fiber with a *g*-ratio around 0.6 has been theoretically shown to provide maximal speed of conducting neural signals in peripheral nerves [Rushton, 1951] but *g*-ratio values

The authors declare no conflict of interest.

Contract grant sponsor: University Medical Center Hamburg-Eppendorf Postdoctoral Fellowship; Contract grant sponsor: Marie Skłodowska-Curie Actions; Contract grant number: MSCA-IF-2015 (EU Horizon 2020)

*Correspondence to: Isabel Ellerbrock, PhD, Department of Systems Neuroscience, University Medical Center Hamburg-

Eppendorf, Martinistr. 52, Building W34, 20246 Hamburg, Germany. E-mail: i.ellerbrock@uke.uni-hamburg.de

Received for publication 25 April 2017; Revised 11 October 2017; Accepted 16 October 2017.

DOI: 10.1002/hbm.23858

Published online 1 November 2017 in Wiley Online Library (wileyonlinelibrary.com).

have been reported to be larger in the CNS in experimental data [Guy et al., 1989; Mason et al., 2001]. It has been suggested that in the healthy condition axons and their microscopic substructures (e.g., their g -ratio) are finely tuned biological devices and that changes of their composition can lead to clinical syndromes [Coggan et al., 2015]. This makes the g -ratio a functionally important property of axons and as such a relevant marker in patients and healthy volunteers alike. Estimating the g -ratio in vivo, however, was not possible until recently. Instead, other white-matter microstructure MRI markers were estimated, such as myelin density or fiber/axon density, for example, in aging [Callaghan et al., 2014; Chang et al., 2015].

Stikov et al. introduced a biophysical model to relate the aggregate g -ratio, that is, the approximation of the g -ratio in a voxel indicating axon properties represented by a summary scalar, to the myelin volume fraction (MVF) and fiber volume fraction (FVF) of a given voxel [Stikov et al., 2011]. The fact that it relates two established in vivo quantitative MRI (qMRI) techniques, namely MVF and FVF mapping, to a functionally relevant microscopic property of axons, namely the g -ratio, makes this model practically relevant for use in neuroscientific and clinical research. This aggregate g -ratio and how it relates back to the microscopic g -ratio was further explored and refined in recent years [West et al., 2016]. In a clinical context, the potential added value of such an in vivo mapping of the MR g -ratio was indicated in a single multiple-sclerosis patient case [Stikov et al., 2015]: the MR g -ratio was highly increased in “new” but not in “old” lesions, indicating that it may provide information regarding re- and demyelination processes. We have recently demonstrated the feasibility of MR g -ratio imaging at the group level [Mohammadi et al., 2015]. Our results were encouraging for clinical and neuroscientific research because the group-level g -ratio (i) was significantly different between fiber pathways, demonstrating its capability to disentangle functionally relevant fiber properties and (ii) was in qualitative agreement with previous ex vivo histology gold standard g -ratio measures for different white matter fiber pathways, indicating its relation to the microscopic g -ratio. However, in vivo MR g -ratio imaging in its current form involves a direct [Cercignani et al., 2017; Mohammadi et al., 2015; Stikov et al., 2015] or indirect [Duval et al., 2017] calibration procedure to relate the MRI markers to the MVF and FVF. More importantly, simulations have demonstrated that the accuracy of the MR g -ratio will depend on the calibration scheme: The more accurate the calibration, the more the MR g -ratio will be weighted by the microscopic g -ratio [Campbell et al., 2017]. To stress this limitation of the MR g -ratio, we will use from now on the term “ g -ratio weighted imaging” as suggested by [Campbell et al., 2017; Duval et al., 2017] instead of the previous “MR g -ratio” notation. The calibration procedure plays a critical role for group comparison studies. For example, in the aging study of Cercignani et al. (2017), the g -ratio was calibrated using

the younger group and the estimated calibration constant was applied to the whole group. However, whether the calibration constant that was determined within one group can be applied to another has not been investigated until now.

Another important conceptual problem that needs to be answered prior to the use of the g -ratio weighted imaging technique in clinical and neuroscientific group studies is which qMRI marker is most favorable for measuring MVF and FVF? Typical qMRI markers for MVF that have been used in the past for g -ratio weighted imaging are (i) magnetization transfer (MT) saturation rate [Campbell et al., 2017; Helms et al., 2009], (ii) quantitative MT imaging (see, e.g., [Campbell et al., 2017; Cercignani et al., 2017]), (iii) macromolecular tissue volume (MTV) based on proton density (PD) imaging [Duval et al., 2017], or (iv) multicomponent relaxometry techniques quantifying the myelin water fraction [MacKay et al., 1994, 2006]. An adaptation of the latter was used in [Dean et al., 2016; Melbourne et al., 2014] for g -ratio weighted imaging. Typical qMRI markers for FVF are based on: (i) Neurite orientation dispersion and density imaging (NODDI) [Zhang et al., 2012] as used, for example, in Stikov et al. [2015] and (ii) tract-fiber density (TFD) [Reisert et al. 2013], as used in Mohammadi et al. [2015], or (iii) the diffusion tensor imaging (DTI)-based fractional anisotropy used in Berman et al. [2017] and Stikov et al. [2011], where the latter measure is valid only in the corpus callosum. One important difference lies in the number of acquired diffusion shells indicating differences in the diffusion effect, which is reflected in scan time that is needed to estimate these imaging markers (e.g., at least two diffusion shells are necessary to estimate the NODDI parameters whereas only one diffusion shell is needed to estimate the TFD metric). A first attempt to compare g -ratio weighted maps based on different imaging markers for MVF was conducted by Campbell et al. [2017]. They demonstrated that equivalent g -ratio weighted maps can be obtained using quantitative MT imaging or the more time-efficient MT saturation approach, whereas the g -ratio based on the clinically popular MT ratio [Tofts, 2004] deviated from the quantitative MT-based g -ratio.

The modeling framework of g -ratio weighted imaging as introduced by Stikov et al. [2011, 2015] is defined for the white matter tissue only. Consequently, partial volume effects in the white matter can bias the g -ratio. This effect is particularly relevant at tissue boundaries, for example, between white matter–gray matter and white matter–cerebrospinal fluid (CSF) boundaries. Importantly, partial volume effects will vary between subjects because of the variation in anatomy and, thus, can be intensified at the group level by residual misalignment of inter-individual white matter tracts. One approach to reduce partial volume effects is to use a mask at the group level, as it is often done in voxel-based morphometry style analyses, which has been recently also applied in g -ratio

weighted imaging [Cercignani et al., 2017; Mohammadi et al., 2015].

In this study, we will build upon the comparative work of Campbell et al. (2017) and test the repeatability and comparability of four different methods for estimating g-ratio weighted maps, one newly introduced method and three previously used approaches [Duval et al., 2017; Mohammadi et al., 2015; Stikov et al., 2015]. The repeatability over time points is tested by assessing the test-retest reliability of g-ratio weighted metrics in a scan-rescan experiment. The comparability is assessed twofold: between g-ratio methods and between two groups of healthy subjects. The first group from the sample cohort is used to estimate the calibration constant and the second as an independent group for comparison. The anticipated variability of the g-ratio values between groups is assessed for each g-ratio method and used to define a lower boundary for the minimum effect size that will be detectable in group studies. Moreover, we test the influence of two key processing steps, namely calibration and masking of unreliable voxels (e.g., due to partial volume effects) at the group level, on repeatability and comparability.

METHODS

Image Acquisition and Subjects

Sample

The study sample consisted of two groups: 12 healthy volunteers (mean 25.6 ± 2.2 years, range 22–30 years, 8 female) were in the first group and 10 healthy volunteers (mean 25.8 ± 2.8 years, range 20–30 years, 3 female) in the second group. All participants were recruited locally at the University Medical Center Hamburg-Eppendorf. They were systematically screened to ensure they were free of any lifetime history of neurological or psychiatric illness, provided written informed consent before being included in the study and received remuneration for participation. The study complied with the principles of the Declaration of Helsinki and was approved by the local ethics committee (Ärztchamber Hamburg). The second group was used solely for the between-group comparison, that is, Figures 8 and 9C.

Data acquisition and software

All MRI sessions were performed on a full-body 3T Tim TRIO scanner (Siemens, Erlangen, Germany) using a 32-channel radiofrequency head coil for receive and body coil for transmission. Foam padding was used to stabilize participants' head and minimize motion. An extensive whole-brain qMRI protocol was acquired consisting of (i) multiparameter mapping (MPM) based on multiecho 3D fast low-angle shot (FLASH) and (ii) diffusion-weighted imaging (DWI). For all participants, MPM and DWI acquisition were conducted twice within one week (6–8 days interval)

to test repeatability of g-ratio measures. Data analyses were performed using MATLAB (The MathWorks, Natick, MA, version R2014b), including customized MATLAB tools, and statistical parametric mapping (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK). All preprocessing steps were deployed identically for both sets of images acquired in two sessions. In imaging MVF and FVF for the purpose of computing g-ratio weighted metrics, all processing steps were performed as suggested in the original publications (see the sections “Estimation of fiber volume fraction (FVF)” and “Estimation of myelin volume fraction (MVF)”).

Diffusion-weighted imaging (DWI). DWI measurements were performed with parallel imaging (GRAPPA, in-plane acceleration factor 2) [Griswold et al., 2002] and simultaneous multislice acquisitions (“multiband,” slice acceleration factor 2) [Feinberg et al., 2010; Moeller et al., 2010; Xu et al., 2013] as described in Setsompop et al. [2012]. The corresponding image reconstruction algorithm was provided by the University of Minnesota Center for Magnetic Resonance Research. A twice-refocused spin echo scheme [Reese et al., 2003] was used to collect multishell DWI composed of 60 diffusion-weighted images acquired with noncollinear diffusion-gradient directions at two b values ($b = 1000$ s/mm² and $b = 2000$ s/mm²), and 12 images with no diffusion-weighting that were interspersed. The total of 132 measurements were repeated twice with identical parameters but reversed phase encoding direction in order to correct for susceptibility-related image distortions to increase the signal-to-noise ratio (SNR). The following parameters were used: 86 slices, slice thickness of 1.6 mm with no gap, 7/8 partial Fourier imaging in phase encoding direction, $224 \times 224 \times 138$ mm² field of view (FOV), resolution 1.6 mm isotropic, echo time (TE) of 122 ms and volume repetition time (TR) of 7.1 s. This resulted in a total acquisition time of ~ 37 min for diffusion-weighted images per subject.

Multiparameter mapping (MPM). To estimate high-resolution magnetization transfer saturation (MT) maps and proton density (PD) maps at 800 μ m the following protocols were acquired: First, rapid calibration data were acquired to correct for inhomogeneities in the RF transmit field [Lutti et al., 2010, 2012]. Second, highly SNR-efficient spoiled multiecho 3D fast low angle shot (FLASH) data were acquired with predominantly PD, T1, or MT weighting according to the MPM protocol (Weiskopf et al., 2013). The flip angle was 60° for the PD- and MT-weighted volumes and 210° for the T1-weighted acquisition. MT weighting was achieved through the application of a Gaussian RF pulse 2 kHz off resonance with 4 ms duration and a nominal flip angle of 220°. To ensure whole-brain coverage at an isotropic resolution of 800 μ m, data were acquired using a FOV of 256 mm head-foot, 224 mm anterior-posterior (AP), and 166 mm right-left (RL). Gradient echoes were acquired with alternating readout gradient

TABLE I. Method and purpose of preprocessing steps of diffusion-weighted images

Method	Purpose
1. Rician bias correction	Correct for Rician signal bias
2. Coregistration	Coregister blip-down to blip-up data
3. Motion correction	Correct for individual motion effects
4. Hyperelastic susceptibility artifact correction (HySCO)	Correct for image distortions due to magnetic field inhomogeneities
5. Diffusion tensor fitting	Produce fractional anisotropy (FA) and average of $b = 0$ maps
6. Creation of brain and white matter mask	Multichannel segmentation of FA and $b = 0$ maps to masks
7. Robust diffusion tensor fitting	Produce maps of outlier-weights for each DWI image
8. Weighted combination of two DWI datasets with reversed phase encoding	Each pair of DWI with reversed phase encoding is combined separately using weights from robust fitting

polarity at equidistant echo times ranging from 2.34 to 18.44 ms in steps of 2.30 ms using a readout bandwidth of 488 Hz/pixel. For PD- and T1-weighted acquisition 8 echoes were acquired, and for the MT-weighted acquisition 6 echoes to maintain a repetition time (TR) of 25 ms for all FLASH volumes. To accelerate the data acquisition, partially parallel imaging using the GRAPPA algorithm was employed in each phase-encoded direction (AP and RL) with 40 reference lines and a speed up factor of two. Total scanning time of the three FLASH datasets was ~ 25 min.

Data processing

Preprocessing of DWI data. Preprocessing of DWI (see summary, Table I) was performed in eight steps using the ACID toolbox (<http://www.diffusiontools.com>): (1) DWIs were corrected for the contribution of Rician bias using the method described in André et al. [2014]. (2) The first $b = 0$ image of each DWI dataset with reversed phase encoding was coregistered using the blip-up data as target and the $b = 0$ image of the blip-down data as source, the resulting transformation was applied to all blip-down data. (3) DWIs were subjected to motion correction using a multitarget registration approach similar to [Mohammadi et al., 2014]. Note that we refrained from eddy current distortion correction, because the employed acquisition protocol included parallel imaging and the twice-refocused spin echo scheme making these distortions negligible. (4) DWIs were corrected for susceptibility distortion artifact (HySCO) using the $b = 0$ images acquired in reversed phase encoding gradients to estimate the fieldmap and apply it to all other images [Ruthotto et al., 2012, 2013]. (5) DWIs were subjected to ordinary least square diffusion tensor fitting to deliver fractional anisotropy (FA) and average of $b = 0$ maps for mask creation. (6) Multichannel segmentation of $b = 0$ and FA images were used to estimate gray, white and CSF tissue segments. Using all three tissue segments a brain mask was created and, based on the white matter segment alone, a white matter mask was calculated. (7) Robust diffusion tensor [Mohammadi et al., 2012, 2013] was used to create maps of outlier weights $\omega_{\uparrow}(i)$ and $\omega_{\downarrow}(i)$ for each DWI image i and blip-up/-down directions \uparrow, \downarrow . (8) Finally, each pair of images of the two

DWI datasets with reversed phase encoding were combined using the robust-fitting weighting maps as follows:

$$wS(i) = \frac{\omega_{\uparrow}(i) S_{\uparrow}(i) + \omega_{\downarrow}(i) S_{\downarrow}(i)}{\omega_{\uparrow}(i) + \omega_{\downarrow}(i)} \quad (1)$$

with $S(i)$ being the diffusion weighted image acquired with the diffusion gradient in the i th direction.

Magnetization transfer (MT) rate and proton density (PD) using the ESTATICS model. All data were coregistered to address inter-scan motion. Maps of $R2^*$ and the PD-, T1-, and MT-weighted maps at $TE = 0$ were estimated from the gradient echoes from all contrasts using the ordinary least squares ESTATICS approach [Weiskopf et al., 2014]. In addition, the image data for each acquired weighting (PDw, T1w, MTw) were averaged over the first six echoes to obtain a second set of PD-, T1-, and MT-weighted maps with increased SNR [Helms and Dechent, 2009]. For PD map calculation, the first set of data (i.e., the PD-weighted images) was approximated at $TE = 0$ to correct for the effect of the $T2^*$ signal decay. Second, the receive field bias was corrected using the UNICORT approach [Weiskopf et al., 2011]. Third, the bias-corrected approximation at $TE = 0$ was converted into a PD map by calibration, assuming that the PD in white matter on average is 69% [Tofts, 2004]. To calculate the MT maps, the second set of data was used because the MT map is not directly affected by $T2^*$ effects and, thus, the dataset with the higher SNR could be used. For both scenarios, the maps were estimated as previously described, resulting in MT and PD maps in percent units [Helms et al., 2008a; Weiskopf et al., 2013]. The MT map depicts the percentage loss of signal (MT saturation) that results from the application of the off-resonance MT prepulse and the dynamics of the magnetization transfer [Helms et al., 2008b]. The PD map depicts the percentage of water molecules in a voxel.

g-Ratio-Weighted Imaging Methods

Estimation of fiber volume fraction (FVF)

Tract fiber density (TFD) as proxy for FVF. For tensor fiber density (TFD) estimation, the Freiburg Fibertools [Reisert et al., 2011, 2013] were used. The TFD in a voxel

was inferred from the relative fiber orientation distribution including the usage of a global scaling factor. Based on the notion that fibers do not terminate within the white matter, this indicator of fiber density aims to establish a global fiber density that takes surrounding voxels into account [Reisert et al., 2013]. First, fiber orientation distribution was computed from the preprocessed DWI data of the first shell ($b = 1000$) using constrained spherical deconvolution within a subject-specific brain mask. DWI data of the first shell only were used to utilize a standard DTI protocol, which may be more readily available from clinically feasible DWI, as used by Reisert et al. [2013]. TFD was then computed from the fiber orientation distribution within an individual white matter mask, obtained from the DWI data (Table I, step 6).

Neurite orientation dispersion and density imaging (NODDI)-based proxy for FVF. For calculation of NODDI parameter maps, the NODDI toolbox [Zhang et al., 2012] was used. The preprocessed data were used to estimate the following NODDI model parameters: the intercellular volume fraction (Aicvf) and the isotropic CSF volume fraction (Aiso). The Aicvf and Aiso compartments were combined with the qMRI markers for MVF (see the section “Estimation of myelin volume fraction (MVF)”) to calculate the FVF according to the relation proposed by Stikov et al. [2015]:

$$FVF = (1 - MVF) (1 - Aiso) Aicvf \quad (2)$$

Estimation of myelin volume fraction (MVF)

MT-based proxy for MVF. The MT saturation maps that were used in this study [Helms et al., 2008b] can be related to the solution of the binary spin-bath model of the MT FLASH sequence derived by Pike [Pike, 1996]. This indicates a correlation between the MT saturation and the transfer term, and, thus, the macromolecular fraction. The sensitivity of the MPM-based MT saturation metric to myelination has been demonstrated in neuroscientific [Callaghan et al., 2014; Dick et al., 2012; Sereno et al., 2013] and clinical research studies [Freund et al., 2013; Grabher et al., 2015].

PD-based proxy for MVF. Mezer et al. [2013] recently proposed a PD-based quantitative imaging technique measuring macromolecular tissue volume (MTV). They defined MTV as the complement of the water volume fraction ($MTV = 1 - PD_{WM}/PD_{CSF}$). The sensitivity of MTV to myelin has been demonstrated through the relationship of MTV with other quantitative myelin mapping techniques [Mezer et al., 2013]. A detailed discussion between MTV and more established imaging markers for the myelin volume can be found in Duval et al. [2017] and a relation to histological estimates for MVF can be found in Berman et al. [2017]. It has been demonstrated that MTV is sensitive to myelin changes over the lifespan and in

demyelinating multiple sclerosis lesions [Yeatman et al., 2014]. The MTV map can be created from our data by using the equation:

$$MTV = 1 - \frac{PD}{100} \quad (3)$$

g-Ratio estimation methods

The g-ratio weighted metrics can be expressed as a function of MTV and FVF:

$$g = \sqrt{1 - \frac{MVF}{FVF}} \quad (4)$$

The derivation of this expression was first described by Stikov et al. [2011] for parallel axons with constant g-ratios. We recently refined its relation to a distribution of microscopic g-ratios [Mohammadi et al., 2015], Stikov et al. [2015] generalized it to an ensemble of microscopic g-ratios with arbitrary orientation, West et al. [2016] further generalized its relation to an ensemble of microscopic g-ratios with an arbitrary radius.

g-Ratio weighted metrics were calculated for each subject and session using different combinations of the qMRI markers for MVF and FVF (see Fig. 1 for an overview). Because the qMRI markers are not measuring absolute values of MVF and FVF but are rather proportional to these, it is necessary to calibrate these [West et al., 2017]. We estimate for each g-ratio weighted metric one distinct calibration constant ($\alpha_1 - \alpha_4$, see the section “Estimation of calibration constants”). Below, the dependence of each g-ratio weighted metric on the calibration constant and the qMRI markers is detailed.

- i. To calculate the first g-ratio metric, g_1 , MT saturation maps [Helms et al., 2008b] were used to determine MVF and TFD [Reisert et al., 2013] for estimating FVF, using the relation $\frac{MVF}{FVF} \sim \frac{\alpha_1 MT}{TFD}$. Hence,

$$g_1 = \sqrt{1 - \frac{\alpha_1 MT}{TFD}} \quad (5)$$

- ii. The second g-ratio metric, g_2 , was estimated using MTV [Mezer et al., 2013] as an estimate for MVF and TFD [Reisert et al., 2013] for FVF, using the relation $\frac{MVF}{FVF} \sim \frac{\alpha_2 MTV}{TFD}$. Hence,

$$g_2 = \sqrt{1 - \frac{\alpha_2 MTV}{TFD}} \quad (6)$$

- iii. The third g-ratio metric, g_3 , was calculated by employing MT maps [Helms et al., 2008b] for determining MVF and NODDI [Stikov et al., 2015] to calculate FVF, using the relation described in Eq. (2). Hence,

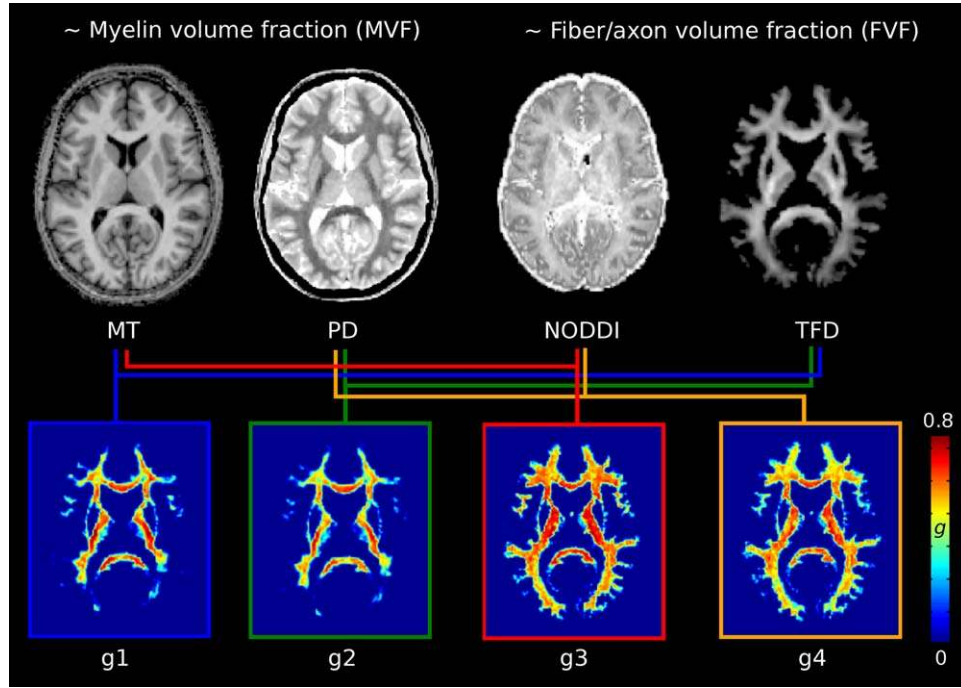


Figure 1.

Four MR *g*-ratio metrics and their dependency on myelin volume fraction (MVF) and fiber volume fraction (FVF). Depicted are (top-left) two quantitative MRI markers for MVF as estimated by the magnetization transfer (MT) saturation map and by the macromolecular tissue volume (MTV) using the proton density (PD) map, (top-right) two imaging markers for FVF as estimated by neurite orientation dispersion and density imaging (NODDI) and tensor

fiber density (TFD), (bottom) the dependency of the four MR *g*-ratio metrics (*g*1–*g*4) on the MVF and FVF imaging markers: *g*1 ~ MT and TFD (blue), *g*2 ~ PD and TFD (green), *g*3 ~ MT and NODDI (red), *g*4 ~ PD and NODDI (orange). Exemplary maps for one subject are displayed for illustration purposes at [x y z; 0 –16 9]. [Color figure can be viewed at wileyonlinelibrary.com]

$$g_3 = \sqrt{1 - \frac{\alpha_3 \text{MT}}{(1 - \alpha_3 \text{MT})(1 - A_{\text{iso}})A_{\text{icvf}}}} \quad (7)$$

- iv. Last, we calculated the fourth *g*-ratio metric, *g*4, by using MTV [Mezer et al., 2013] as proxy for MVF and metrics derived from the NODDI model [Stikov et al., 2015] for FVF, using the relation described in Eq. (2). Hence,

$$g_4 = \sqrt{1 - \frac{\alpha_4 \text{MTV}}{(1 - \alpha_4 \text{MTV})(1 - A_{\text{iso}})A_{\text{icvf}}}} \quad (8)$$

Note that *g*2 is a novel approach to estimate the *g*-ratio weighted map, whereas *g*1 is using the same approach as in Mohammadi et al. [2015]. Assuming that MT and quantitative MT maps are equally valid proxies for MVF as proposed by Campbell et al. [2017], *g*4 differs from Duval et al. [2017] only by the additional calibration constant and *g*3 is similar to the approach used by Stikov et al. [2015].

Estimation of calibration constants

The calibration constant in each *g*-ratio was determined by minimizing Eqs. (5–8) using a literature value for the *g*-ratio (denoted “*g*-ratio calibration value”) within the corpus callosum splenium of the sample average (see ROI 8 in Fig.4 in [Mohammadi et al., 2015]). Instead of performing the calibration on a single subject as in Mohammadi et al. [2015], we used the sample-averaged imaging markers for myelin and fiber density in group space (see the section “Spatial alignment” for details) to estimate the calibration constants. This group-based calibration was used to reduce interindividual variation in the calibration procedure. The resulting calibration constants (Table II) were used to calculate the *g*-ratio map for each subject within both groups. Note that the calibration was based on the first group only.

To test the influence of the calibration constant on the repeatability and *g*-ratio values of different *g*-ratio weighted imaging methods, we used three different *g*-ratio calibration values that were reported in the literature for humans: (i) *g* = 0.6 [Rushton, 1951], (ii) *g* = 0.7 [Graf von

TABLE II. Calibration constants for g-ratio weighted metrics

Calibration value	α_1	α_2	α_3	α_4
0.6	0.0928	0.5542	0.1731	1.0333
0.7	0.0740	0.4416	0.1478	0.8824
0.8	0.0522	0.3118	0.1137	0.6790

Calibration constants (α_1 – α_4) for each g-ratio weighted metric (g_1 – g_4) as determined in the calibration step using three reported literature values of $g = 0.6$ [Rushton, 1951], $g = 0.7$ [Graf von Keyserlingk and Schramm, 1984], and $g = 0.8$ [Wharton and Bowtell, 2012].

Keyserlingk and Schramm, 1984], and (iii) $g = 0.8$ [Wharton and Bowtell, 2012]. The results are reported in Figure 3. For the other analyses, we used a g-ratio value of $g = 0.7$, because this value was closest to what is known from ex vivo histology. It was based on human histological evaluation of microscopic g-ratios in humans for large-diameter axons, which are frequent in the callosal splenium [Graf von Keyserlingk and Schramm, 1984]. Of note, the calibration constant was determined individually for each g-ratio metric. The g-ratio weighted metrics were calculated only within the white matter using the MT-based white matter (c2) tissue probability maps (TPM) thresholded at $c2 > 0.95$. Physically meaningless g-ratio values (e.g., with imaginary values), and g-ratios based on meaningless MVF and FVF estimates (i.e., $A_{icvf} \leq 0$, $MT \leq 0$, $MTV \leq 0$, $TFD \leq 0$ and $FVF \leq 0$), were set to zero.

Spatial Alignment, Processing, and Analysis

Spatial alignment

For the scan–rescan analysis, all data acquired per subject were registered (i) across the two imaging methods, DWI and MPM and (ii) across the two time points, using a rigid-body transformation (spm_coreg). First, for each time point separately, MPM images were registered to the DWI, using the PD-weighted image as source image and the $b = 0$ image as target and applying the estimated transformation to the MT and PD map. Second, MPM and DWI data from the second time point were registered to the data from the first time point, using the skull-stripped MT map from second time point as source and the skull-stripped MT map from the first time point as target, and applying the estimated transformation to all DWI and MPM data from the second time point.

Then, the data were spatially normalized to MNI space using the DARTEL SPM software [Ashburner, 2007]. To this end, individual median MT maps (across time points) were produced and segmented into gray- and white-matter probability maps [Ashburner and Friston, 2005]. Then, these MT white-matter and gray-matter segments were used as input in DARTEL to estimate the deformation fields using an existing template from a previous study [Mohammadi et al., 2015]. The estimated

deformation fields were applied to MPM and DWI maps, and to all g-ratio maps (for details of computation of g-ratio maps, see the section “g-Ratio estimation methods”).

Masking of unreliable voxels using tissue probability map (TPM)-based masks

White-matter tissue probability maps (TPM) of subjects in group space were averaged and thresholded to create the TPM-based masks. TPMs estimate partial volume information based on tissue classes [Ashburner and Friston, 2005]. They contain probabilistic information about white matter brain tissue, ranging from 0 (no white matter) to 1 (only white matter). In a first analysis (Fig. 3), the optimal threshold of the TPM mask was determined by sweeping the threshold from 0 to 1. The optimal threshold was identified as the point at which the rate of increase in repeatability was minimal. For the other analyses, a liberal white matter mask was created using a threshold of 0 (Fig. 4A) and a conservative white matter mask was created using the data-driven optimal threshold from Figure 3 (Fig. 4B). The liberal white matter mask was used as reference (denoted “without additional masking”) and the conservative white matter mask was used for masking of unreliable voxels (e.g., due to partial volume effects) at the group level (denoted as “TPM mask”).

Fiber tracts as regions of interest (ROIs)

The group average g-ratios were calculated within specific fiber tracts in MNI space, as defined in the white matter atlas of the Anatomy toolbox [Eickhoff et al., 2005]: callosal body (cb), corticospinal tracts (ct), cingulum (cing), inferior occipitofrontal fasciculus (iof), optic radiation (or), and superior longitudinal fasciculus (slf) (Fig. 2). Each fiber tract-specific ROI was generated by thresholding the probabilistic fiber tract maps at 50%, including an additional masking step by either the white matter or TPM mask. Furthermore, an additional voxel threshold was applied to ensure that only ROIs that contain more than 15 voxels were used in the analysis. Only the fiber tract ROIs were included in the analysis, for which neither mean nor the absolute difference in g-ratio values between time points contained invalid data points (NaNs) in at least $n = 8$ of the subjects in each group.

Repeatability of g-ratio-weighted metrics in predetermined fiber tracts

The repeatability of g-ratio weighted methods was tested using within-subject scan–rescan measurements:

$$rg = 100 \times \frac{1}{N} \sum_{i=1}^N \left(1 - \frac{\Delta g(i)}{g_{cal}} \right), \text{ and } \Delta g(i) = \frac{1}{N_{ROI}} \sum_{r=1}^{N_{ROI}} |g_{t0}(r) - g_{t1}(r)|, \quad (9)$$

with r being the index of the r th voxel and N_{ROI} the number of voxels in the ROI, $\Delta g(i)$ being the absolute

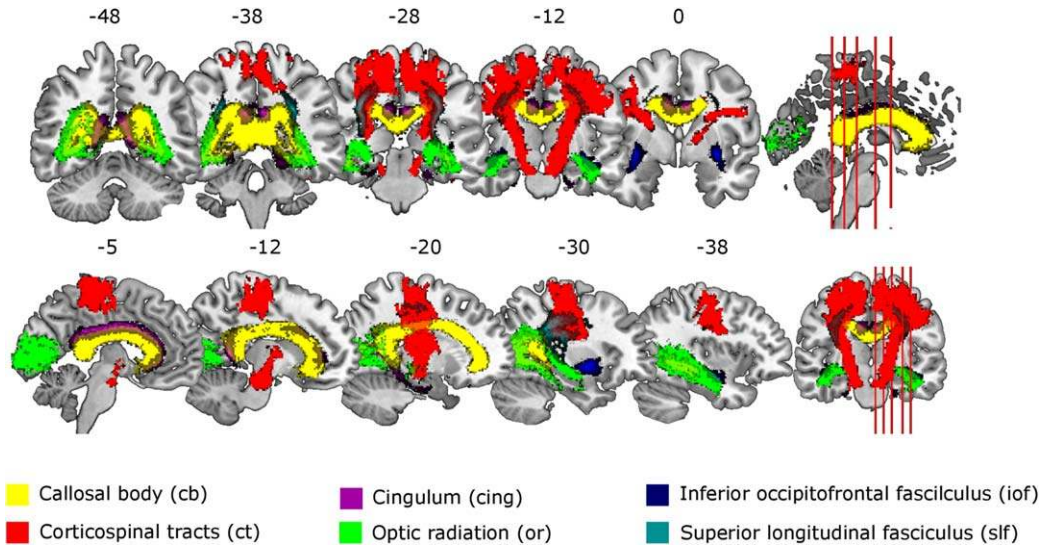


Figure 2.

Regions of interest for tract-specific analyses. Fiber tracts were selected from the SPM anatomy toolbox (Eickhoff et al., 2005) to perform tract-specific analyses. [Color figure can be viewed at wileyonlinelibrary.com]

difference between g -ratio maps of two time points (t_0 and t_1) averaged over the ROI, g_{cal} being the g -ratio calibration value, and i being the index of the i th subject and N the number of subjects in group 1. The g -ratio difference maps were normalized by the g -ratio calibration value to ensure that the repeatability does not depend on the mean of g -ratios within each g -ratio weighted imaging method. In order to avoid skewed values in the difference g -ratio weighted maps, only g -ratio values above zero were included into the analysis. Results are reported on a 0–100 scale, ranging from no repeatability (0%) to perfect repeatability (100%).

Comparability of g -ratio-weighted methods between methods and groups

First, the comparability between the four g -ratio weighted methods was assessed by comparing mean g -ratio values in predetermined fiber tract ROIs. Second, Pearson's correlation coefficient r was calculated across fiber tract ROIs for each combined pair of g -ratio methods (six combinations). Note that for the correlation analysis, the mean of the g -ratio for each subject was used separately, that is, the correlation coefficient was estimated from 12×5 data points (see the section “Quantifying between-subject, between-tract, and between-group comparability” for information on the included fiber tract ROIs). The comparability between g -ratio weighted maps were calculated for both groups. To this end, the group specific g -ratio mean was calculated for each fiber tract ROI and g -ratio method.

Quantifying between-subject, between-tract, and between-group comparability

The variability between groups (group), tracts (tract), and subjects (sub) were assessed using the following root-mean-square error (RMSE) measures:

$$RMSE_{group} = \sqrt{\frac{1}{N} \sum_{j=1}^N (g_{k1,j} - g_{k2,j})^2} \quad (10)$$

$$RMSE_{tract} = \sqrt{\frac{1}{N} \sum_{j=1}^N (g_{k1,j} - \langle g_{k1} \rangle_{ROI})^2} \quad (11)$$

$$RMSE_{sub} = \frac{1}{N} \sum_{j=1}^N RMSE_{sub,j}, \text{ and} \quad (12)$$

$$RMSE_{sub,j} = \sqrt{\frac{1}{M} \sum_{s=1}^M (g_{k1,j,s} - \langle g_{k1,j} \rangle_{group})^2}$$

with j being the index of the j th fiber tract, N the number of fiber tracts, s being the index of the s th subject, M the number of subjects, $k1$ and $k2$ being group 1 and 2, $\langle g_{k1} \rangle_{ROI}$ the mean of group-averaged g -ratios across ROIs, and $\langle g_{k1,j} \rangle_{group}$ the mean of subject-specific g -ratios within each ROI. Note that in this analysis the inferior occipitofrontal fasciculus (iof) tract was omitted, that is, resulting in $N = 5$ ROIs, because there were not enough voxels in the ROI after TPM masking.

RESULTS

The Calibration Procedure

The literature g value used in the calibration process (g -ratio calibration value) was found to influence g -ratios and

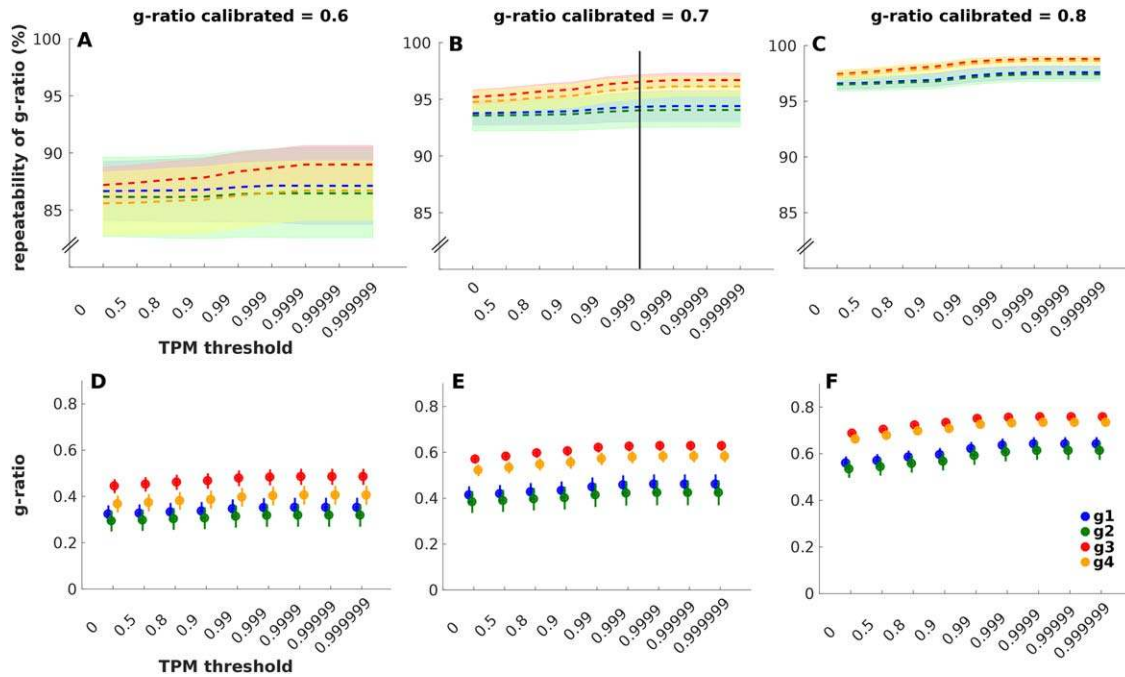


Figure 3.

The g -ratio values (bottom) and their repeatability (top) averaged across fiber tracts as a function of the threshold of the tissue probability maps (TPM threshold) using three different calibration values: $g = 0.6$ (A, D), $g = 0.7$ (B, E), and $g = 0.8$ (C, F). On the y-axis, the repeatability is depicted, ranging from 0 to 100. On the x-axis, the threshold for generating the TPM mask is depicted, ranging from 0 to 1: voxels with a tissue probability

smaller than the threshold were excluded. Error shades and error bars represent the standard deviation. The optimal threshold (0.999, black line in B) for the TPM mask was determined as the value after which the repeatability decreased by $<15\%$ (illustrated in Fig. 4B). The four g -ratio methods are depicted in blue (g1), green (g2), red (g3), and orange (g4). [Color figure can be viewed at wileyonlinelibrary.com]

repeatability across g -ratio weighted methods. An increase in the g -ratio calibration value led to (i) an increase in g -ratios across metrics, (ii) an increase in repeatability, (iii) a reduction in the between-method variation of the repeatability, and (iv) a reduction in standard deviation (Fig. 3). While differences in repeatability were particularly pronounced between TFD-based and NODDI-based methods for the higher g -ratio calibration values ($g = 0.7$ and $g = 0.8$), there was no clear tendency for the lowest g -ratio calibration value ($g = 0.6$). In general, the calibration constants (α_1 – α_4 , Table II) tended to decrease with increasing g -ratio calibration value. Interestingly, the calibration constant for the method 4 (α_4) was close to 1 for the lowest g -ratio calibration value ($g = 0.6$), which corresponds to applying no calibration constant.

Number of Voxels in Each Fiber Tract

The number of voxels in each fiber tract for each g -ratio weighted methods and each masking approach are reported in Table III. The number of voxels in the inferior occipitofrontal fasciculus (iof) was below the threshold of 15 voxels when the TPM masking approach was applied, suggesting

high interindividual variation in white matter tissue probability or high misalignment between subjects in this tract. Overall, the number of voxels was higher in the NODDI-based g -ratio methods than in the TFD-based methods.

Repeatability of g -Ratio-Weighted Methods

Repeatability of g -ratio weighted maps without additional masking to reduce unreliable voxels at the group level (mask in Fig. 4A) displayed differences between g -ratios weighted methods (Fig. 6A). The lowest value of 78% was found in the inferior occipitofrontal fasciculus (iof) using method g2, and the highest value of 93% was found in the corticospinal tracts (ct) using method g3. In general, the NODDI-based measures showed higher repeatability across fiber tracts (g3: $\sim 86\%$, g4: $\sim 85\%$) compared to TFD-based metrics (g1: $\sim 82\%$, g2: $\sim 82\%$): the influence of the imaging marker for FVF on the repeatability was about $\sim 4\%$ on average. In comparison, the imaging marker for MVF changed the repeatability on average only about $\sim 1\%$.

Using the TPM mask (mask in Fig. 4B) to reduce the number of unreliable voxels (e.g., due to partial volume

TABLE III. Average number of voxels across Group 1 in each tract with and without TPM masking approach

Tract	Nvox g1	Nvox g2	Nvox g3	Nvox g4
<i>w/o</i>				
<i>cb</i>	11486	11074	13053	13005
<i>ct</i>	5402	4978	6274	6265
<i>cing</i>	496	440	936	933
<i>iof</i>	232	204	243	241
<i>or</i>	2976	2386	5772	5741
<i>slf</i>	391	384	409	409
<i>TPM</i>				
<i>cb</i>	6735	6555	7329	7328
<i>ct</i>	3600	3475	3677	3677
<i>cing</i>	64	61	86	86
<i>iof</i>	0*	0*	0*	0*
<i>or</i>	1973	1684	2815	2813
<i>slf</i>	355	350	365	365

Note that only tracts with voxel sizes above the threshold (>15) were included in the analyses; tracts with no voxels (marked with an asterisk (*)) were excluded from the analyses. (Nvox) number of voxels, (w/o) without additional masking, (TPM) tissue probability maps, (cb) callosal body, (ct) corticospinal tracts, (cing) cingulum, (iof) inferior occipitofrontal fasciculus, (or) optic radiation, (slf) superior longitudinal fasciculus.

effects), repeatability was slightly elevated across fiber tracts (range across tracts: 80–93%, Fig. 6B). TFD-based metrics displayed overall lower repeatability (84% and 83%) than NODDI-based measures (90% and 89%, Fig. 6A,B). Importantly, the difference in repeatability between the g-ratio methods using different FVF imaging markers (higher in NODDI than TFD) was increased to $\sim 7\%$ when applying the TPM mask as compared to no correction, whereas the difference between g-ratios methods using different MVF imaging markers remained $\sim 1\%$, suggesting that the NODDI-based g-ratios profit strongest from the TPM masking approach. Note that in Figure 6B, the g-ratios were not presented within the inferior occipitofrontal fasciculus (iof) because the number of voxels in this mask was below the threshold (Table III).

Comparability Between g-Ratio-Weighted Methods

The variation across tracts was most pronounced for the TFD-based methods. The lowest value was found in the optic radiation using g-ratio method 2 ($g_2 = 0.25$), and the highest in the corticospinal tract using g-ratio method 3 ($g_3 = 0.66$). The same trend detected in the repeatability was observed for the mean g-ratios: the NODDI-based methods (g_3, g_4) showed higher g-ratio values across fiber tracts as compared to TFD-based methods (g_1, g_2). On average the mean g-ratio was increased by about 21% with respect to the g-ratio calibration value when changing the imaging marker for FVF ($g_3 - g_1 = 0.16$ and $g_4 - g_2 = 0.14$),

whereas changing the imaging marker for MVF only led to a change of about 6% ($g_1 - g_2 = 0.03$ and $g_3 - g_4 = 0.05$). The two imaging markers for the MVF showed the most prominent difference ($\sim 10\%$) in the corticospinal tract ($g_2 - g_1 = 0.07$ and $g_4 - g_3 = 0.08$, Fig. 6C). Along this tract, a higher variation was visible in the MTV-based g-ratio maps than in the MT-based g-ratio maps (Fig. 5).

After TPM masking, all g-ratios increased slightly (mean g-ratios: g_1 from 0.41 to 0.46, g_2 from 0.38 to 0.42, g_3 from 0.57 to 0.63, and g_4 from 0.52 to 0.58, Fig. 6D). Importantly, after masking, the across-tract variability decreased more strongly for the NODDI-based g-ratios ($\sim 50\%$) than for the TFD-based g-ratios ($\sim 20\%$) (Fig. 9A). Interestingly, the variability between subjects appeared to be independent of the masking approach (Fig. 9B).

All combinations of g-ratio methods showed significant correlations when using the masking approach (Fig. 7). However, g-ratio methods with the same FVF imaging marker (g_1 & g_2 and g_3 & g_4 ; Fig. 7B) displayed higher correlation coefficients with smaller confidence intervals than g-ratio methods using different FVF imaging markers.

Comparability of g-Ratios between Groups

The g-ratio values across tracts were compared between different groups when applying the calibration constants derived from group 1 to both groups (Fig. 8). In line with previous observations on repeatability and between-method comparability, the NODDI-based g-ratio methods showed smaller intergroup variability than the TFD-based g-ratio methods (Fig. 9C). TPM masking reduced the between-group variability for the NODDI-based methods (g_3, g_4), whereas it even increased the between-group variability for the TFD-based methods (g_1, g_2). Using TPM masking, the variability between groups was smallest for g-ratio method 3 (3.5%), and largest for method 2 (14%). g-Ratio methods 1 and 4 displayed a variability of 11% and 4% between groups, respectively (Fig. 9C).

Note that applying the TPM masking led to missing values in the inferior occipitofrontal fasciculus (iof) in both groups (Fig. 8B), which is why we removed it from the quantitative analysis (Fig. 9), that is, all presented measures were calculated across the five remaining tracts.

DISCUSSION

In this study, we used two acquisition methods (EPI-based diffusion MRI and FLASH-based multi parameter mapping, MPM) to estimate four distinct MRI markers for the MVF and FVF. Based on these MRI markers four equally legitimate methods for g-ratio weighted imaging were compared with respect to their repeatability and comparability. In addition, the effect of the calibration and masking of unreliable voxels (e.g., due to partial volume effects) at the group level was tested. Our main findings were: (i) overall repeatability and comparability depended

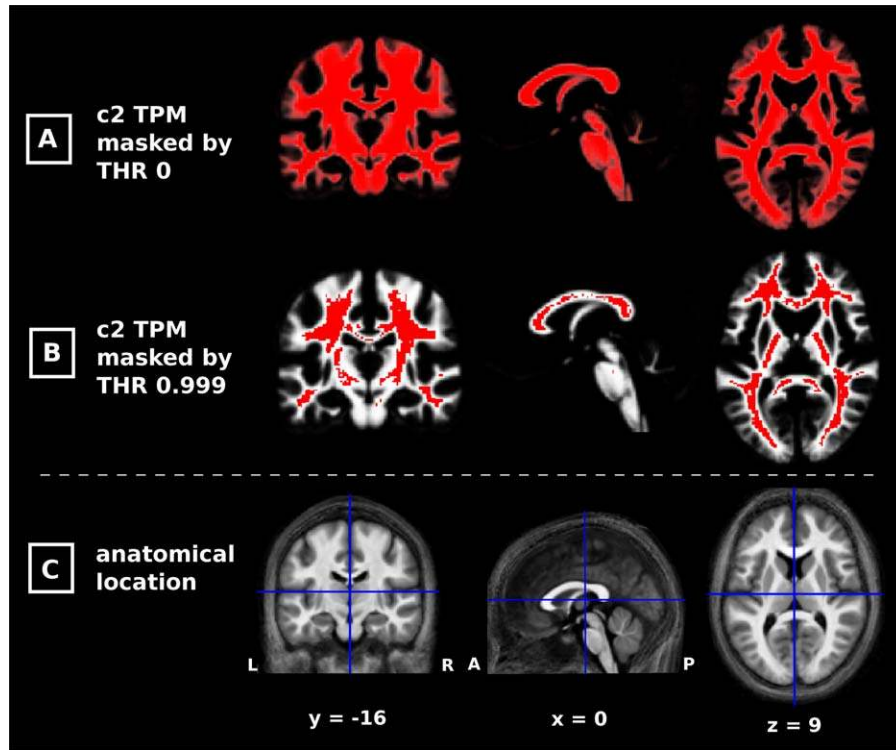


Figure 4.

White matter tissue probability map (TPM) and the corresponding masks (overlaid in red). The mean white matter (c2) TPM is overlaid by two masks: **(A)** a threshold of 0 is used (we refer to this as “without additional masking”) and **(B)** the optimal threshold (THR = 0.999, for details see Fig. 3) is used (we refer

to this as “TPM mask”). **(C)** Group averaged MT images are displayed for anatomical orientation. All images are displayed at $[x \ y \ z; 0 \ -16 \ 9]$. (THR) threshold, (L) left, (R) right, (A) anterior, and (P) posterior. [Color figure can be viewed at wileyonlinelibrary.com]

largely on the method used to estimate the FVF (NODDI outperformed TFD), (ii) repeatability and comparability could be improved when masking out unreliable voxels, (iii) the g -ratio method based on NODDI and MT showed the highest repeatability (90%) and lowest variability between groups (3.5%), (iv) the calibration step was crucial: calibrating to a lower g -ratio calibration value (i.e., $g = 0.6$ instead of the commonly used $g = 0.7$) not only affected the repeatability and comparability, but also the reported dependence on the FVF imaging marker.

The MRI Markers

The repeatability of different g -ratio weighted imaging methods and their comparability depended highly on the choice of the FVF imaging marker (i.e., NODDI vs TFD). The mean g -ratio and its repeatability was higher for the NODDI-based than for the TFD-based g -ratio methods (the g -ratio mean about 21%, and its repeatability by about 4%, Fig. 6). Moreover, the comparability between g -ratio methods was highest if the same imaging marker for FVF was used (Fig. 7). Interestingly, the correlation between the

TFD-based methods was higher than the correlation between NODDI-based methods. This was probably caused by the higher dynamic range of TFD-based g -ratios as compared to the NODDI-based g -ratios. Both FVF imaging markers were based on the same diffusion-weighted imaging (DWI) technique but the underlying computational approach and DWI protocol were different: the NODDI-based approach required multishell DWI data to solve a biophysical multicompartiment signal model including a compartment of the axonal volume fraction, whereas the TFD-based approach required single-shell DWI data to estimate the FVF from the orientation distribution of fiber pathways, as typically used in tractography. The most important advantage of the NODDI model over the TFD-based approach is the generative signal model that directly relates the diffusion signal to microstructure compartments. However, in this model, specific parameters are fixed using prior knowledge to stabilize the problem. These model assumptions, which are controversially discussed in the literature [Dhital et al., 2015; Jelescu et al., 2016], can be a reason for the observed variation of fitted NODDI parameters between 1.5 T and 3 T [Chung et al., 2016]. The advantage of the TFD approach compared to

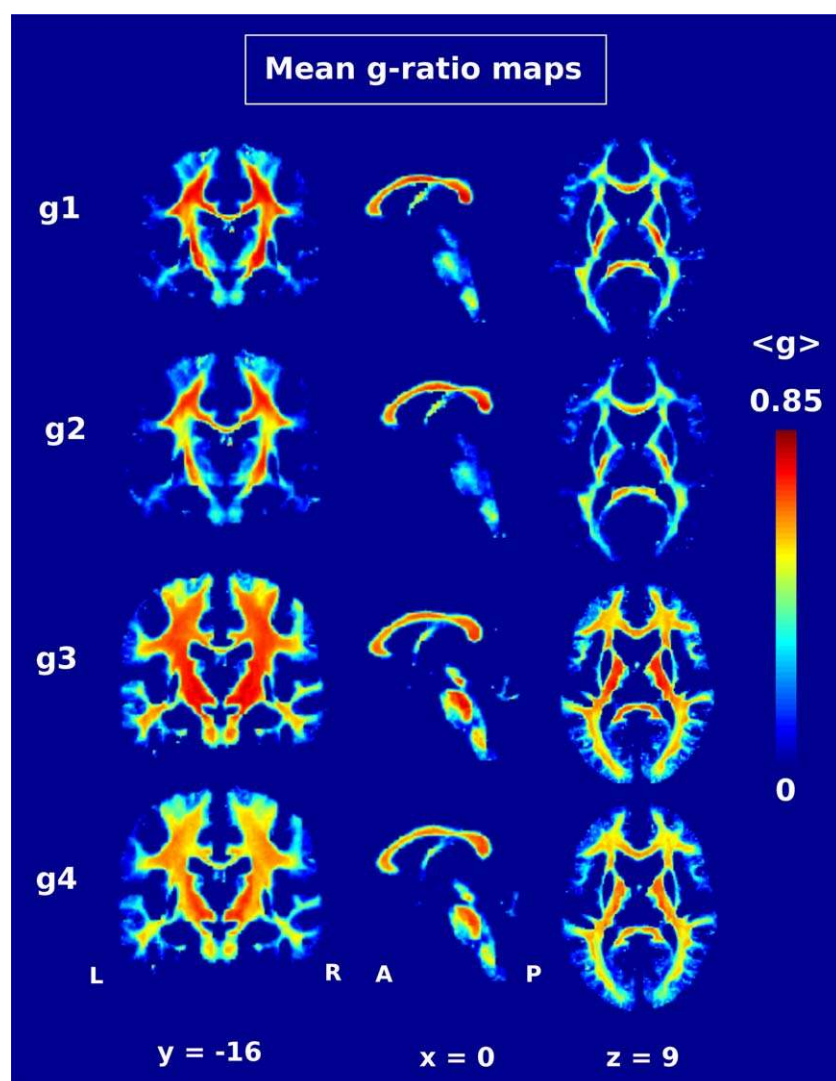


Figure 5.

Averaged *g*-ratio maps for each *g*-ratio-weighted method (*g*1–*g*4). Most prominently, an apparent thinning of white matter tracts in the *g*-ratio maps were observed for the TFD-based (*g*1 and *g*2) as compared to the NODDI-based (*g*3 and *g*4) *g*-ratio methods. Although less pronounced, some local deviations, for example, along the corticospinal tracts (compare left column),

were also apparent if the imaging marker for the myelin compartment was varied, that is, between MT-based (*g*1 and *g*3) and MTV-based (*g*2 and *g*4) methods. All images are displayed at [*x y z*; 0 –16 9]. (L) left, (R) right, (A) anterior, and (P) posterior. [Color figure can be viewed at wileyonlinelibrary.com]

NODDI is that it is more flexible concerning the DWI acquisition protocol, making it more suitable for clinical application, for example, by reducing the acquisition time or making it compatible with single-shell tractography DWI data (e.g., HARDI data). One obvious disadvantage of TFD compared to NODDI is that single-shell data is used to estimate the FVF, which encodes less microstructural information than two-shell data. Another important disadvantage of the TFD approach will be discussed in the section “*The masking approach at the group level.*”

Although less pronounced, the choice of imaging marker for MVF also influenced the repeatability (~1% difference, Fig. 6A) and the mean *g*-ratio (~6% difference, Fig. 6C). Across all tracts, *g*-ratio values were higher for the MT-based (*g*1, *g*3) than for the MTV-based *g*-ratio methods (*g*2, *g*4). Even though both metrics are related to the myelin compartments, they use fundamentally different MRI techniques: MTV uses PD mapping and is sensitive to membranes and proteins [Mezer et al. 2013], whereas MT uses an off-resonance pulse to probe the

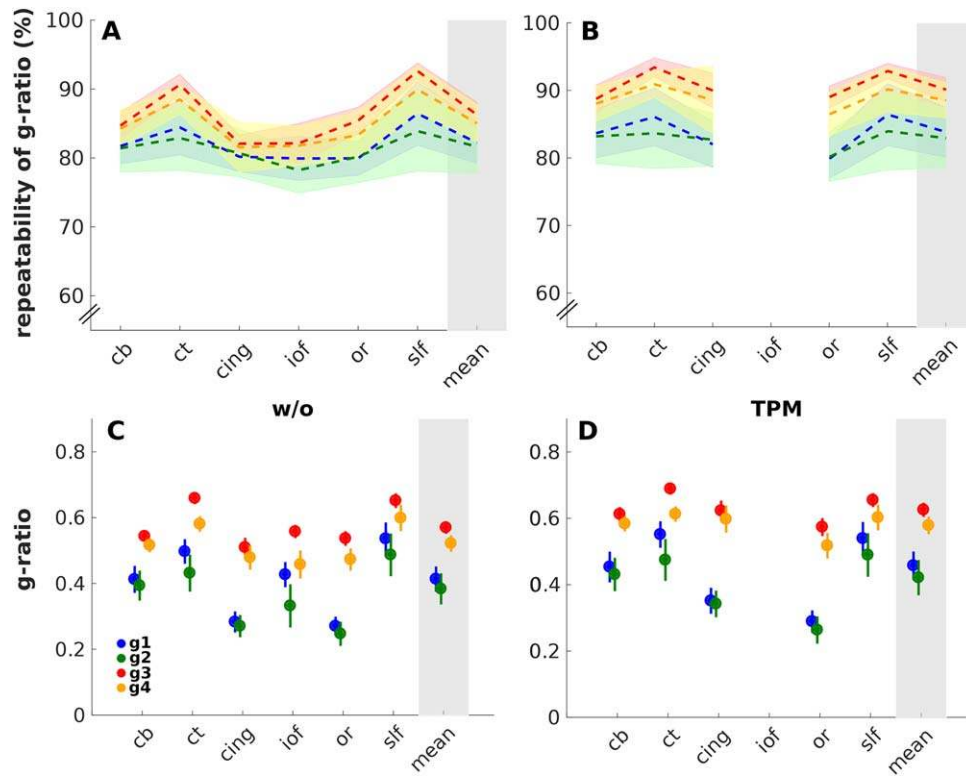


Figure 6.

Comparison of repeatability and mean g -ratios between four g -ratio-weighted methods ($g1$ – $g4$) in predefined fiber tracts (Fig. 2). **(A, B)** Repeatability (Eq. 9) of g -ratio methods and its standard deviation across subjects (shaded areas): **(A)** without additional masking (**w/o**) (Fig. 4A) **(B)** and with use of the TPM mask (Fig. 4B). **(C, D)** g -Ratio values and their standard deviation across subjects (error bars): **(C)** and **(D)** use the approaches described in **(A)** and **(B)**. Without additional masking and after TPM masking, the g -ratio methods differed mostly between

TFD- and NODDI-based methods. Note that after TPM masking, there were no voxels in the *iof* tract and, thus, no value is presented in **(B)** and **(D)** for the *iof*. (*cb*) callosal body, (*ct*) corticospinal tracts, (*cing*) cingulum, (*iof*) inferior occipitofrontal fasciculus, (*or*) optic radiation, (*slf*) superior longitudinal fasciculus. The four g -ratio methods are depicted in blue ($g1$), green ($g2$), red ($g3$), and orange ($g4$). [Color figure can be viewed at wileyonlinelibrary.com]

macromolecular fraction of the tissue [Callaghan et al., 2015; Helms et al., 2008b]. Our data (Figs. 5 and 6C) support the notion that these imaging markers are sensitive to similar but also distinct microscopic information. The MTV-based g -ratio weighted maps ($g2$, $g4$) showed more heterogeneity across the white matter than the MT-based g -ratio weighted maps ($g1$, $g3$). The difference between MVF imaging markers was particularly strong in the corticospinal tract. While both imaging markers have been previously used as proxy for myelin density (MT [Mohammadi et al., 2015], MTV [Duval et al., 2017]), there is only direct comparison with ex vivo histology gold standard for MTV [Berman et al., 2017].

The Masking Approach at the Group Level

We introduced a new masking method to reduce partial volume effects at the group level, namely the TPM-based masking approach. Our data showed that masking increased

the repeatability and g -ratio values, and decreased the between-group variability in the NODDI-based g -ratio methods almost by the same amount, indicating that partial volume effects are the main source for variation in the NODDI-based g -ratio methods. We hypothesize that the TFD-based g -ratio methods were less affected by using the TPM masking approach than the NODDI-based methods because partial volume effects are not the only relevant source of variation in the TFD method. The method to estimate the TFD uses the local vicinity of the voxel [Reisert et al., 2013]. As a result, adjacent voxels at tissue boundaries that are not white matter can bias the TFD values in the white matter. On the contrary, NODDI works on a voxel-by-voxel basis only (i.e., it is a local method). In other words, the TFD-based g -ratios not only are affected by partial volume effects in the raw data but also by additional mixture of tissue compartments at the modelling level (e.g., when mixing signal from adjacent voxels at tissue boundaries).

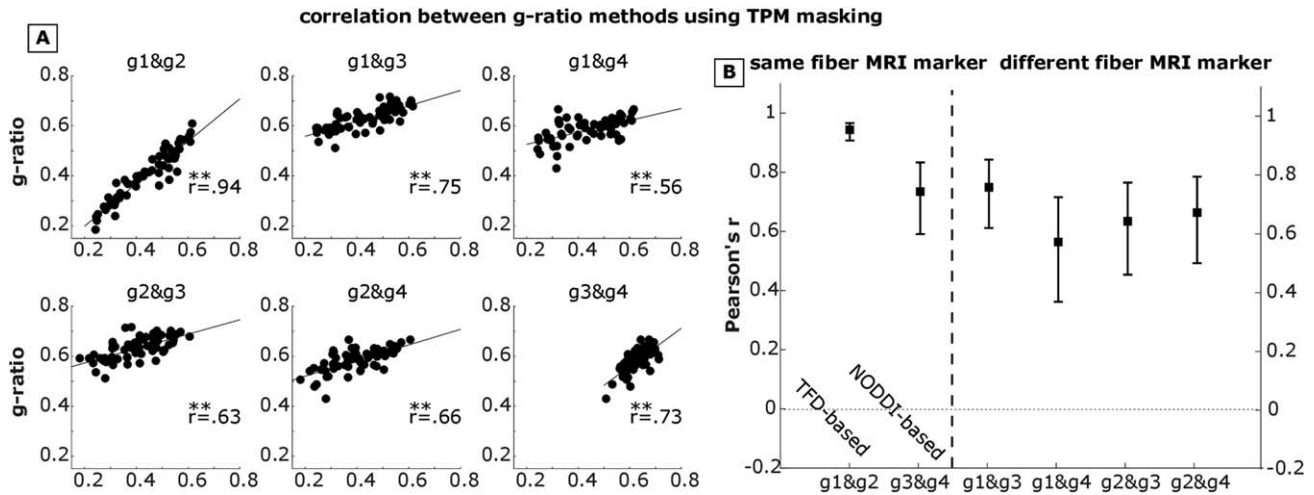


Figure 7.

Correlation between g -ratio methods in predetermined white matter tracts. **(A)** Displayed are g -ratio values and correlation coefficients (Pearson's r) between all combinations of the four g -ratio weighted methods ($g1$ – $g4$) for each subject (black dots) in six predefined white matter pathways (Fig. 2) using TPM masking. **(B)** Depicted are the respective Pearson's r with 95% confidence interval. On the left of the dashed line, pairs of g -ratio methods using the same fiber

imaging marker (either TFD or NODDI) are depicted. On the right, pairs of g -ratio methods with different fiber imaging marker are displayed. All g -ratio combinations correlated significantly, however, Pearson's r between g -ratios was strongest and the confidence interval was smaller (or comparable) when the same fiber imaging marker is used. $**P < 0.01$.

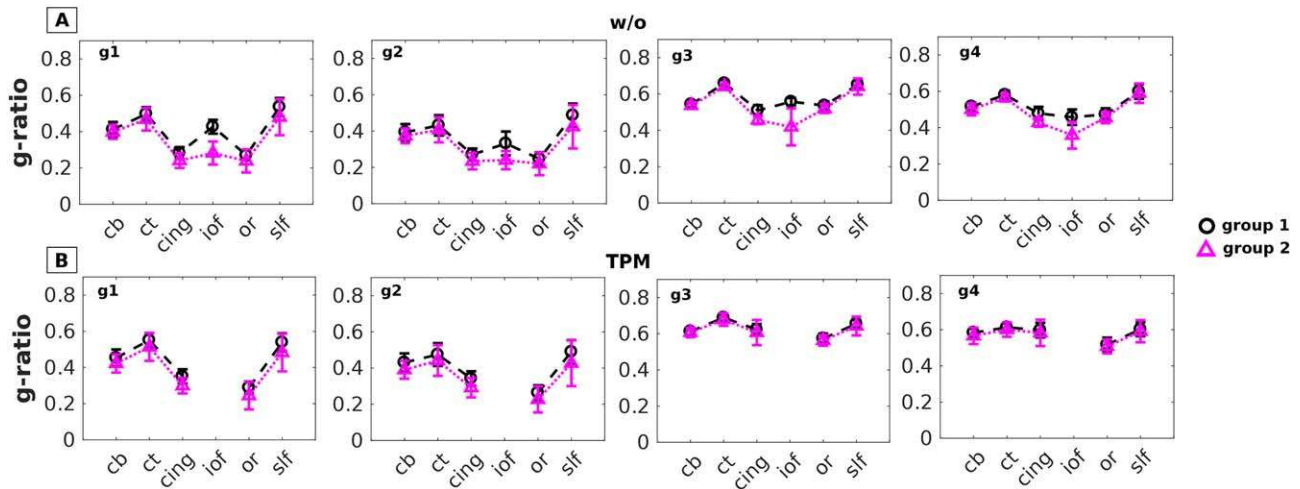


Figure 8.

Comparability of g -ratio values between groups for each g -ratio method. Displayed are g -ratios of the group 1 (black) and the group 2 (magenta) using **(A)** no additional masking, and **(B)** TPM masking. g -Ratio values are comparable between groups. Quantification of the across group variability is depicted in Figure 9C. Note that TPM masking resulted in missing values in

the inferior occipitofrontal fasciculus (iof). Error bars represent the standard deviation. (TPM) tissue probability maps, (w/o) without additional masking, (cb) callosal body, (ct) corticospinal tracts, (cing) cingulum, (or) optic radiation, (slf) superior longitudinal fasciculus. [Color figure can be viewed at wileyonlinelibrary.com]

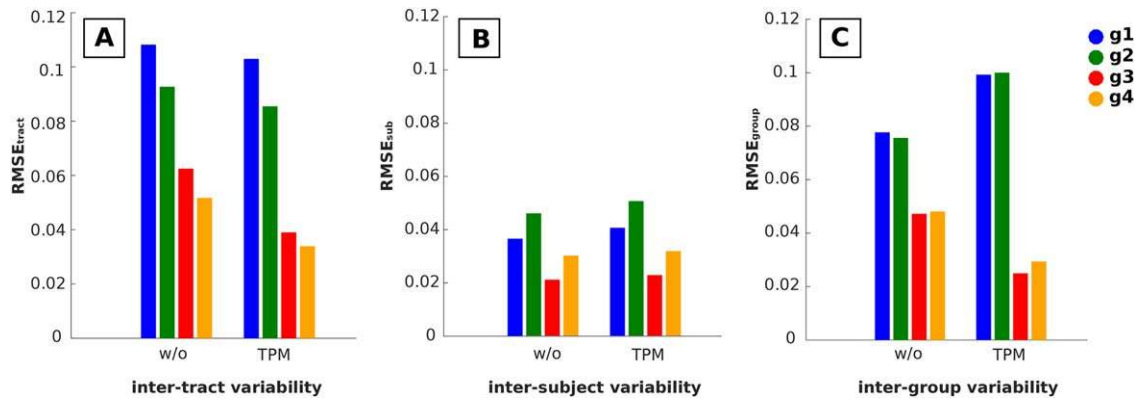


Figure 9.

Variability in g -ratio across tracts (A), subjects (B), and groups (C) using four g -ratio weighted methods. The root mean square error (RMSE) was used to assess (A) variability across tracts between white matter fiber tracts (Eq. 11), (B) variability between subjects (Eq. 12), and (C) variability across groups (Eq. 10). The variability across tracts and groups was strongly

affected when using TPM masking. More specifically, TPM masking reduced the variability of the NODDI methods ($g3$, $g4$) but had little effect on the TFD methods. The four g -ratio methods are depicted in blue ($g1$), green ($g2$), red ($g3$), and orange ($g4$). (TPM) tissue probability maps, (w/o) without additional masking. [Color figure can be viewed at wileyonlinelibrary.com]

Despite being potentially more reasonably available due to reduced acquisition time, TFD-based g -ratio methods cannot be recommended for applicability in neuroscientific or clinical studies due to their reduced repeatability and comparability between groups.

The Calibration Procedure

The calibration step is currently the most critical processing step in g -ratio weighed imaging [Campbell et al., 2017]. Various calibration values for the g -ratio in humans have been used (g between 0.6 and 0.8, e.g., in Graf von Keyserlingk and Schramm, 1984; Rushton, 1951; Wharton and Bowtell, 2012) to calibrate the MRI markers to MVF (and FVF). However, it is unclear which one is the most accurate one and should be, therefore, used for calibration. Here, we showed that the g -ratio and its repeatability across methods increased (as expected) with increasing g -ratio calibration values. Whether the sensitivity to potential group differences will be reduced with increasing g -ratio calibration value is beyond the scope of this study but an interesting question that could be investigated in future studies. Using higher g -ratio calibration values ($g = 0.7$, $g = 0.8$), the choice of the FVF imaging marker was the crucial factor that determined the repeatability and g -ratio values: NODDI-based g -ratios and their repeatability were higher than TFD-based g -ratios. Interestingly, the dependency of the repeatability on the MVF and FVF imaging markers was more complicated if the lowest g -ratio calibration value ($g = 0.6$) was used: almost no difference in repeatability between MVF imaging markers for the TFD-based g -ratios but the largest difference in

repeatability between MVF imaging markers for the NODDI-based g -ratios were observed (Fig. 3).

In general, the calibration constants, that is, the constant by which the MRI markers are scaled to relate them to MVF and FVF, tended to decrease with increased g -ratio calibration values. Thereby, the calibration constant for the g -ratio method 4 became almost one ($\alpha = 1.0333$) when using the lowest g -ratio calibration value. This means that using no calibration constant (i.e., $\alpha = 1$), as proposed by Duval et al. (2017), would correspond to using the calibration scheme proposed here with a g -ratio calibration value of approximately $g = 0.6$. Note, however, even in Duval and colleagues' (2017) approach, where there is no calibration constant necessary to relate MTV to MVF, there is an indirect calibration step involved, namely when estimating the PD maps.

Determining the calibration constant in a control group and applying it to another, as done in Cercignani et al. [2017], seems to be a reasonable approach for g -ratio group studies. However, it is unclear whether the same calibration constant can be applied to different groups without influencing the group results. Using the same calibration constant (derived from group 1) on two different groups, we observed the smallest inter-group variability (3.5%) for g -ratio method 3 after TPM masking (Fig. 9C), that is, when using NODDI and MT as FVF and MVF imaging markers. Conclusively, group differences larger than 4–5% might be detectable when using the g -ratio method 3.

Methodological Considerations

The calibration constant: The calibration constant does not only vary for different methods (Table II) but probably also when using different MRI protocols and might even

be specific to the MRI system (e.g., 3 T vs 1.5 T). Although it was out of the scope of this study to test the dependency of the calibration constant with respect to these factors, this aspect will be important in future research. We recommend estimating the calibration constant individually for each MRI protocol and MR system based on a healthy group of subjects.

Signal-to-noise ratio (SNR) in DWI: The SNR of the $b = 0$ image was on average about 66. Assuming a mono-exponential signal decay the SNR for a diffusion weighted image with $b = 2000 \text{ s/mm}^2$ can be estimated to be about $9.3 = 66 \cdot \exp(-2)$. Usually, an SNR value of at least 10 is recommended for diffusion weighted images (see e.g., Jones et al. [2013]). To correct for potential Rician bias in the diffusion signal, we used the method recommend by André et al. (2014).

Note that the total scan time for the diffusion MRI protocol could be reduced to about 19 minutes, therefore halving the acquisition time used in this study, if only one phase encoding direction and a single-refocused spin echo scheme was acquired. This leads to a reduction of the total SNR by about 24% when accounting for reduced SNR ($\sim 29\%$) due to halving the number of images and increased SNR ($\sim 5\%$) due to reduced echo time by using single-refocused spin echo.

Fiber tract masks: In most of the tracts, there were more than 50 voxels after applying TPM masking. In the inferior occipitofrontal fasciculus (iof), however, the number of voxels could get below 15 voxels (compare Table III). A possible explanation for the high variability is that this particular tract has a small spatial extent (Fig. 2) and probably showed high inter-individual variability. To maintain comparability, we removed the iof tract from the quantitative analyses in Figure 9.

Threshold for TPM masking: The repeatability and mean g-ratio increased, when a higher threshold (up to the threshold of $\text{THR} = 0.999$) was used for TPM masking. However, along with an increasing threshold, the number of voxels within each fiber tract decreased. To find a balance between optimal partial volume correction and maximal number of voxels in each tract, we chose the smallest threshold at which the increase in repeatability was less than 15% (see Fig. 3, $\text{THR} = 0.999$).

Underestimation of g-ratio values: The g-ratio values reported in this study were consistently below the calibration g-ratio value of $g = 0.7$. While a potential reason for the extremely low g-ratio values using TFD (Fig. 6) has been discussed above, one reason why NODDI-based g-ratio values were also consistently below the calibration value could be related to the fact that we used independent masks for calibrating the g-ratio and for the analyses. The calibration was based on a hand-drawn mask (similar to ROI 8, fig. 4 in Mohammadi et al. [2015]), including only the central part of the splenium of the corpus callosum, where the g-ratio is particularly high. The mean g-ratio values in white matter tracts were calculated based on the Jülich atlas for white matter [Eickhoff et al., 2005].

Other means to reduce partial volume effects: In our previous study [Mohammadi et al., 2015], an additional adaptive de-noising step during preprocessing of the DWI was performed. Although in the present study we did not include this noise reduction step to maintain comparability with other studies [Cercignani et al., 2017; Duval et al., 2017; Stikov et al., 2015], we believe that for future studies adaptive smoothing might demonstrate an additional opportunity to reduce partial volume effects at the single-subject level because it is designed to preserve (and sharpen) anatomical borders, e.g., between white and gray matter [Becker et al., 2014; Tabelow et al., 2015]. Another approach to reduce partial volume effects in the g-ratio maps could be skeletonization toward the tract center, which has been previously proposed for FA maps [Smith et al., 2006].

Accuracy of MRI markers: Note that in the proposed g-ratio weighted imaging models [Eqs. (5–8)], we assume a linear relation between the MRI markers for MVF and FVF and the true MVF and FVF, respectively. This assumption has only been tested for the MVF [Berman et al., 2017]. To our knowledge, less is known about the relation between MRI markers for FVF and the corresponding histological markers. A linear relation has been reported between diffusion MRI markers for fiber density and ex vivo histology markers using, respectively, Bielschowsky and Luxol fast blue staining [Wang et al., 2013] and myelin staining [Sepehrband et al., 2015].

Note also that more advanced models for MVF estimation have been used for g-ratio mapping, such as quantitative MT [Cercignani et al., 2017; Stikov et al., 2015] or myelin water fraction mapping based on multicompartment T2 mapping [Dean et al., 2016; Melbourne et al., 2014]. It was beyond the scope of our study to test the comparability and repeatability of these latter MRI markers for MVF, but this important issue should be investigated in future studies.

CONCLUSIONS

In conclusion, the g-ratio method 3 (NODDI as a proxy for FVF and MT as a proxy for MVF) performed best in terms of repeatability and comparability between groups. However, it requires the most time consuming MRI protocol (two-shell DWI for NODDI and an extensive MPM protocol for MT).

ACKNOWLEDGEMENTS

The authors are grateful to Jürgen Finsterbusch for support on the MR side, to the University of Minnesota Center for Magnetic Resonance Research for providing the image reconstruction algorithm for the simultaneous multislice acquisitions, and to Martina Callaghan and Nikolaus Weiskopf for valuable discussions and for providing the multiparameter mapping sequences.

REFERENCES

- André ED, Grinberg F, Farrher E, Maximov II, Shah NJ, Meyer C, Jaspar M, Muto V, Phillips C, Balteau E, Baron J-C (2014): Influence of noise correction on intra- and inter-subject variability of quantitative metrics in diffusion kurtosis imaging. *PLoS ONE* 9:e94531. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3983191/>.
- Ashburner J (2007): A fast diffeomorphic image registration algorithm. *NeuroImage* 38:95–113.
- Ashburner J, Friston KJ (2005): Unified segmentation. *NeuroImage* 26:839–851.
- Becker SMA, Tabelow K, Mohammadi S, Weiskopf N, Polzehl J (2014): Adaptive smoothing of multi-shell diffusion weighted magnetic resonance data by msPOAS. *NeuroImage* 95: 90–105.
- Berman S, West KL, Does MD, Yeatman JD, Mezer AA (2017): Evaluating g-ratio weighted changes in the corpus callosum as a function of age and sex. *NeuroImage* <http://www.sciencedirect.com/science/article/pii/S1053811917305475>.
- Callaghan MF, Freund P, Draganski B, Anderson E, Cappelletti M, Chowdhury R, Diedrichsen J, FitzGerald THB, Smittenaar P, Helms G, Lutti A, Weiskopf N (2014): Widespread age-related differences in the human brain microstructure revealed by quantitative magnetic resonance imaging. *Neurobiol Aging* 35:1862–1872.
- Callaghan MF, Helms G, Lutti A, Mohammadi S, Weiskopf N (2015): A general linear relaxometry model of R1 using imaging data. *Magn Reson Med* 73:1309–1314.
- Campbell JSW, Leppert IR, Narayanan S, Boudreau M, Duval T, Cohen-Adad J, Pike GB, Stikov N (2017): Promise and pitfalls of g-ratio estimation with MRI. *NeuroImage* (in press).
- Cercignani M, Giulietti G, Dowell NG, Gabel M, Broad R, Leigh PN, Harrison NA, Bozzali M (2017): Characterizing axonal myelination within the healthy population: a tract-by-tract mapping of effects of age and gender on the fiber g-ratio. *Neurobiol Aging* 49:109–118.
- Chang YS, Owen JP, Pojman NJ, Thieu T, Bukshpun P, Wakahiro MLJ, Berman JL, Roberts TPL, Nagarajan SS, Sherr EH, Mukherjee P, Gong G (2015): White matter changes of neurite density and fiber orientation dispersion during human brain maturation. *PLoS ONE* 10:e0123656.
- Chung AW, Seunarine KK, Clark CA (2016): NODDI reproducibility and variability with magnetic field strength: A comparison between 1.5 T and 3 T. *Hum Brain Mapp* 37:4550–4565.
- Coggan JS, Bittner S, Stiefel KM, Meuth SG, Prescott SA (2015): Physiological dynamics in demyelinating diseases: Unraveling complex relationships through computer modeling. *Int J Mol Sci* 16:21215–21236.
- Dean DC, O'Muircheartaigh J, Dirks H, Travers BG, Adluru N, Alexander AL, Deoni SCL (2016): Mapping an index of the myelin g-ratio in infants using magnetic resonance imaging. *NeuroImage* 132:225–237.
- Dhital I, Kellner B, Reiser M, Kiselev V (2015): Isotropic Diffusion Weighting Provides Insight on Diffusion Compartments in Human Brain White Matter In vivo. In: *Proc Intl Soc Magn Reson Med*. Vol. 23 (Toronto, ON). abstract 2788.
- Dick F, Tierney AT, Lutti A, Josephs O, Sereno MI, Weiskopf N (2012): In vivo functional and myeloarchitectonic mapping of human primary auditory areas. *J Neurosci* 32:16095–16105.
- Duval T, Le Vy S, Stikov N, Campbell J, Mezer A, Witzel T, Keil B, Smith V, Wald LL, Klawiter E, Cohen-Adad J (2017): g-Ratio weighted imaging of the human spinal cord in vivo. *NeuroImage* 145:11–23.
- Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005): A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage* 25:1325–1335.
- Feinberg DA, Moeller S, Smith SM, Auerbach E, Ramanna S, Gunther M, Glasser MF, Miller KL, Ugurbil K, Yacoub E (2010): Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS ONE* 5:e15710.
- Freund P, Weiskopf N, Ashburner J, Wolf K, Sutter R, Altmann DR, Friston K, Thompson A, Curt A (2013): MRI investigation of the sensorimotor cortex and the corticospinal tract after acute spinal cord injury: a prospective longitudinal study. *Lancet Neurol* 12:873–881.
- Grabner P, Callaghan MF, Ashburner J, Weiskopf N, Thompson AJ, Curt A, Freund P (2015): Tracking sensory system atrophy and outcome prediction in spinal cord injury. *Ann Neurol* 78:751–761.
- Graf von Keyserlingk D, Schramm U (1984): Diameter of axons and thickness of myelin sheaths of the pyramidal tract fibres in the adult human medullary pyramid. *Anat Anz* 157:97–111.
- Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A (2002): Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med* 47: 1202–1210.
- Guy J, Ellis EA, Kelley K, Hope GM (1989): Spectra of G ratio, myelin sheath thickness, and axon and fiber diameter in the guinea pig optic nerve. *J Comp Neurol* 287:446–454.
- Helms G, Dathe H, Dechent P (2008a): Quantitative FLASH MRI at 3T using a rational approximation of the Ernst equation. *Magn Reson Med* 59:667–672.
- Helms G, Dathe H, Kallenberg K, Dechent P (2008b): High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D FLASH MRI. *Magn Reson Med* 60:1396–1407.
- Helms G, Dechent P (2009): Increased SNR and reduced distortions by averaging multiple gradient echo signals in 3D FLASH imaging of the human brain at 3T. *J Magn Reson Imaging* 29:198–204.
- Helms G, Draganski B, Frackowiak R, Ashburner J, Weiskopf N (2009): Improved segmentation of deep brain grey matter structures using magnetization transfer (MT) parameter maps. *NeuroImage* 47:194–198.
- Jelescu IO, Veraart J, Fieremans E, Novikov DS (2016): Degeneracy in model parameter estimation for multi-compartmental diffusion in neuronal tissue. *NMR Biomed* 29:33–47.
- Jones DK, Knösche TR, Turner R (2013): White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *NeuroImage* 73:239–254.
- Lutti A, Hutton C, Finsterbusch J, Helms G, Weiskopf N (2010): Optimization and validation of methods for mapping of the radiofrequency transmit field at 3T. *Magn Reson Med* 64:229–238.
- Lutti A, Stadler J, Josephs O, Windischberger C, Speck O, Bernarding J, Hutton C, Weiskopf N (2012): Robust and fast whole brain mapping of the RF transmit field B1 at 7T. *PLoS ONE* 7:e32379.
- MacKay A, Whittall K, Adler J, Li D, Paty D, Graeb D (1994): In vivo visualization of myelin water in brain by magnetic resonance. *Magn Reson Med* 31:673–677.
- MacKay A, Laule C, Vavasour I, Bjarnason T, Kolind S, Mädlar B (2006): Insights into brain microstructure from the T2 distribution. *Magn Reson Imaging* 24:515–525.

- Mason JL, Langaman C, Morell P, Suzuki K, Matsushima GK (2001): Episodic demyelination and subsequent remyelination within the murine central nervous system: changes in axonal calibre. *Neuropathol Appl Neurobiol* 27:50–58.
- Melbourne A, Eaton-Rosen Z, Vita ED, Bainbridge A, Cardoso MJ, Price D, Cady E, Kendall GS, Robertson NJ, Marlow N, Ourselin S (2014): Multi-modal Measurement of the Myelin-to-Axon Diameter g-ratio in Preterm-born Neonates and Adult Controls. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014*. Springer, Cham. pp 268–275. https://link.springer.com/chapter/10.1007/978-3-319-10470-6_34.
- Mezer A, Yeatman JD, Stikov N, Kay KN, Cho N-J, Dougherty RF, Perry ML, Parvizi J, Hua LH, Butts-Pauly K, Wandell BA (2013): Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat Med* 19:1667–1672.
- Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, Uğurbil K (2010): Multiband multislice GE-EPI at 7 Tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med* 63:1144–1153.
- Mohammadi S, Carey D, Dick F, Diedrichsen J, Sereno MI, Reiser M, Callaghan MF, Weiskopf N (2015): Whole-brain in-vivo measurements of the axonal g-ratio in a group of 37 healthy volunteers. *Front Neurosci* 9:441.
- Mohammadi S, Hutton C, Nagy Z, Josephs O, Weiskopf N (2013): Retrospective correction of physiological noise in DTI using an extended tensor model and peripheral measurements. *Magn Reson Med* 70:358–369.
- Mohammadi S, Nagy Z, Hutton C, Josephs O, Weiskopf N (2012): Correction of vibration artifacts in DTI using phase-encoding reversal (COVIPER). *Magn Reson Med* 68:882–889.
- Mohammadi S, Tabelow K, Ruthotto L, Feiweier T, Polzehl J, Weiskopf N (2014): High-resolution diffusion kurtosis imaging at 3T enabled by advanced post-processing. *Front Neurosci* 8:427.
- Pike GB (1996): Pulsed magnetization transfer contrast in gradient echo imaging: a two-pool analytic description of signal response. *Magn Reson Med* 36:95–103.
- Reese TG, Heid O, Weisskoff RM, Wedeen VJ (2003): Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magn Reson Med* 49:177–182.
- Reiser M, Mader I, Anastasopoulos C, Weigel M, Schnell S, Kiselev V (2011): Global fiber reconstruction becomes practical. *NeuroImage* 54:955–962.
- Reiser M, Mader I, Umarova R, Maier S, Tebartz van Elst L, Kiselev VG (2013): Fiber density estimation from single q-shell diffusion imaging by tensor divergence. *NeuroImage* 77:166–176.
- Rushton W. a H (1951): A theory of the effects of fibre size in medullated nerve. *J Physiol (Lond)* 115:101–122.
- Ruthotto L, Kugel H, Olesch J, Fischer B, Modersitzki J, Burger M, Wolters CH (2012): Diffeomorphic susceptibility artifact correction of diffusion-weighted magnetic resonance images. *Phys Med Biol* 57:5715.
- Ruthotto L, Mohammadi S, Heck C, Modersitzki J, Weiskopf N (2013): Hyperelastic susceptibility artifact correction of DTI in SPM. In: *Meinzer, H-P, Deserno, TM, Handels, H, Tolxdorff, T, editors. Bildverarbeitung für die Medizin 2013*. Springer Berlin Heidelberg. Informatik aktuell pp 344–349.
- Sepehrband F, Clark KA, Ullmann JFP, Kurniawan ND, Leange G, Reutens DC, Yang Z (2015): Brain tissue compartment density estimated using diffusion-weighted MRI yields tissue parameters consistent with histology. *Hum Brain Mapp* 36:3687–3702.
- Sereno MI, Lutti A, Weiskopf N, Dick F (2013): Mapping the human cortical surface by combining quantitative T(1) with retinotopy. *Cereb Cortex* 23:2261–2268.
- Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL (2012): Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med* 67:1210–1224.
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TEJ (2006): Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* 31:1487–1505.
- Stikov N, Campbell JSW, Stroh T, Lavelée M, Frey S, Novek J, Nuara S, Ho M-K, Bedell BJ, Dougherty RF, Leppert IR, Boudreau M, Narayanan S, Duval T, Cohen-Adad J, Picard P-A, Gasecka A, Côté D, Pike GB (2015): In vivo histology of the myelin g-ratio with magnetic resonance imaging. *NeuroImage* 4:397–405.
- Stikov N, Perry LM, Mezer A, Rykhlevskaia E, Wandell BA, Pauly JM, Dougherty RF (2011): Bound pool fractions complement diffusion measures to describe white matter micro and macro-structure. *NeuroImage* 54:1112–1121.
- Tabelow K, Mohammadi S, Weiskopf N, Polzehl J (2015): POAS4SPM: a toolbox for SPM to denoise diffusion MRI data. *Neuroinformatics* 13:19–29.
- Tofts P (2004): *Quantitative MRI of the Brain: Measuring Changes Caused by Disease*, 1st ed. John Wiley & Sons.
- Wang S, Chopp M, Nazem-Zadeh M-R, Ding G, Nejad-Davarani SP, Qu C, Lu M, Li L, Davoodi-Bojd E, Hu J, Li Q, Mahmood A, Jiang Q (2013): Comparison of neurite density measured by MRI and histology after TBI. *PLoS ONE* 8:e63511.
- Waxman SG (1980): Determinants of conduction velocity in myelinated nerve fibers. *Muscle Nerve* 3:141–150.
- Weiskopf N, Callaghan MF, Josephs O, Lutti A, Mohammadi S (2014): Estimating the apparent transverse relaxation time (R_2^*) from images with different contrasts (ESTATICS) reduces motion artifacts. *Front Neurosci* 8:278.
- Weiskopf N, Lutti A, Helms G, Novak M, Ashburner J, Hutton C (2011): Unified segmentation based correction of R1 brain maps for RF transmit field inhomogeneities (UNICORT). *NeuroImage* 54:2116–2124.
- Weiskopf N, Suckling J, Williams G, Correia MM, Inkster B, Tait R, Ooi C, Bullmore ET, Lutti A (2013): Quantitative multi-parameter mapping of R1, PD^* , MT, and R_2^* at 3T: a multi-center validation. *Front Neurosci* 7.
- West KL, Kelm ND, Carson RP, Does MD (2016): A revised model for estimating g-ratio from MRI. *NeuroImage* 125:1155–1158.
- West KL, Kelm ND, Carson RP, Gochberg DF, Ess KC, Does MD (2017): Myelin volume fraction imaging with MRI. *NeuroImage* doi: 10.1016/j.neuroimage.2016.12.067. [Epub ahead of print].
- Wharton S, Bowtell R (2012): Fiber orientation-dependent white matter contrast in gradient echo MRI. *PNAS* 109:18559–18564.
- Xu J, Moeller S, Auerbach EJ, Strupp J, Smith SM, Feinberg DA, Yacoub E, Uğurbil K (2013): Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *NeuroImage* 83:991–1001.
- Yeatman JD, Wandell BA, Mezer AA (2014): Lifespan maturation and degeneration of human brain white matter. *Nat Commun* 5:4932.
- Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC (2012): NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* 61:1000–1016.