



Peripheral Nerve Diffusion Tensor Imaging

Interreader and Test-retest Reliability as Quantified by the Standard Error of Measurement

Fabian Preisner¹ · Philipp Bäumer^{1,2} · Michaela Wehrstein³ · Birgit Friedmann-Bette³ · Matthes Hackbusch⁴ · Sabine Heiland¹ · Martin Bendszus¹ · Moritz Kronlage¹

Received: 26 September 2019 / Accepted: 16 November 2019 / Published online: 5 December 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract

Purpose Diffusion tensor imaging (DTI) is increasingly being used in magnetic resonance neurography (MRN). The purpose of this study was to determine the interreader and test-retest reliability of peripheral nerve DTI in MRN with focus on the sciatic nerve.

Methods In this prospective study 27 healthy volunteers each underwent 3 scans of a short DTI protocol on separate days consisting of a T2-weighted turbo spin-echo and single-shot DTI sequence of the sciatic nerve of the dominant leg. The DTI parameters fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) were obtained after manual nerve segmentation by two independent readers. Intraclass correlation coefficients (ICC), standard error of measurement (SEM), and Bland-Altman plots were calculated as measures for both interreader and test-retest agreement for all readout parameters.

Results The mean ± standard deviation was 0.507 ± 0.05 for FA, $1308.5 \pm 162.4 \times 10^{-6} \text{ mm}^2/\text{s}$ for MD, $905.6 \pm 145.4 \times 10^{-6} \text{ mm}^2/\text{s}$ for RD and $2114.1 \pm 219.2 \times 10^{-6} \text{ mm}^2/\text{s}$ for AD. The SEM for FA was 0.02 for interreader and test-retest agreement, the SEM for MD, RD, and AD ranged between $46.2 \times 10^{-6} \text{ mm}^2/\text{s}$ (RD) and $70.1 \times 10^{-6} \text{ mm}^2/\text{s}$ (AD) for interreader reliability and between $45.9 \times 10^{-6} \text{ mm}^2/\text{s}$ (RD) and $70.1 \times 10^{-6} \text{ mm}^2/\text{s}$ (AD) for test-retest reliability. The ICC for interreader reliability of DTI parameters ranged between 0.81 and 0.92 and ICC for test-retest reliability between 0.76 and 0.91.

Conclusion Peripheral nerve DTI of the sciatic nerve is reliable and reproducible. The measures presented here may serve as first orientation values of measurement accuracy when interpreting parameters of sciatic nerve DTI.

Keywords Diffusion tensor imaging · Magnetic resonance imaging · Peripheral nervous system · Reference values · Reproducibility of results

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00062-019-00859-0>) contains supplementary material, which is available to authorized users.

✉ Moritz Kronlage
Moritz.Kronlage@med.uni-heidelberg.de

¹ Department of Neuroradiology, Heidelberg University Hospital, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

² Center for Radiology Dia.log, Vinzenz-von-Paul Str. 8, 84503 Altötting, Germany

³ Department of Sports Medicine (Internal Medicine VII), Medical Clinic, Heidelberg University Hospital, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany

⁴ Institute of Medical Biometry and Informatics, University of Heidelberg, Im Neuenheimer Feld 130.3, 69120 Heidelberg, Germany

Abbreviations

AD	Axial diffusivity
DTI	Diffusion tensor imaging
FA	Fractional anisotropy
ICC	Intra-class correlation coefficient
MD	Mean diffusivity
MRN	Magnetic resonance neurography
RD	Radial diffusivity
SEM	Standard error of measurement

Introduction

Magnetic resonance (MR) imaging with diffusion tensor imaging (DTI) quantifies proton diffusion based on a 3-dimensional ellipsoid model [1, 2]. Since water diffusion is

avored along nerve fiber tracts and hindered perpendicular to the nerve axis by cell membranes and myelin sheaths, peripheral nerve DTI has the potential to deliver specific biomarkers of microstructural nerve integrity and has been investigated in numerous recent studies in healthy volunteers and specific neuropathies [3–11].

The main readout parameters of DTI are fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD) [12]. The FA is a dimensionless value between 0 (equal diffusion in all directions) and 1 (diffusion in only one direction), is considered to be a marker of nerve tissue damage and is reduced in various pathologies [3, 8, 10, 11]. While MD expresses an average of diffusion in all directions, AD and RD specifically describe diffusion along and orthogonal to the nerve course and might possibly offer specific information about the myelin sheath and axonal integrity but are also modulated by inflammation and edema [3, 12–14].

Peripheral nerve DTI is regarded as a promising technique to deliver biomarkers for differential diagnosis and follow-up of various neuropathies. Despite being investigated in various studies [3–11], a broad implementation of peripheral nerve DTI in the clinical routine is still lacking. A fundamental question that has to be addressed before implementation is how reliable the readout parameters are, both between different readers (interreader reliability) and between different scanning sessions (test-retest reliability).

The authors are aware of only a few studies that have addressed the reliability of peripheral nerve DTI and amongst them there is great variety in terms of imaging technique, number of scans, anatomical regions, and postprocessing methods [15–22]. Most of these studies expressed the results by reporting the intraclass correlation coefficient (ICC), a frequently reported dimensionless parameter in reliability studies; however, care should be taken when interpreting absolute ICC values since ICC not only depends on the precision of the test but also on the variance of the assessed parameter in the examined population and is therefore sensitive to data range [23, 24]. In simple terms, this means that a particular test has a high ICC (suggesting an accurate test) in a group where the measured parameter differs largely between individuals but a low ICC (suggesting an inaccurate test) in a group where the measured parameter occurs only in a low range, even if the test is exactly the same. A statistical parameter which describes the precision of the method independently of the variance in the population is the standard error of measurement (SEM) [25]. In addition, SEM not only enables classification of reliability in qualitative terms, such as good or moderate but also represents a quantitative measure of precision, expressed in the same physical unit as the measured quantity [25]. The concept of the SEM is easy to understand in clinical practice when imagining that a certain test

measures values around a true value with a certain error. Assuming a normal distribution, approximately 68% of all measurements the test delivers are ± 1 SEM around the true value and around 95% of all measurements the test delivers are distributed ± 2 SEM around the true value [25]. To our knowledge, a systematic assessment of both interobserver and test-retest SEM is still lacking for all major parameters of peripheral nerve diffusion tensor imaging.

Although advances in imaging techniques now even enable an investigation of even small-caliber nerves, this study still focused on the sciatic nerve as one of the most commonly investigated nerves in MRN, especially when describing systemic neuropathies [5, 26–30]. The relatively straight course and the large caliber of the sciatic nerve represent good technical conditions for MRN, especially for quantitative imaging.

The aim of this study was to determine the interreader and test-retest reliability of peripheral nerve DTI by providing the SEM and ICC of scalar DTI parameters. A cohort of 27 healthy participants were prospectively examined where each received 3 scans of a standardized peripheral nerve DTI protocol, which were analyzed by two independent readers.

Material and Methods

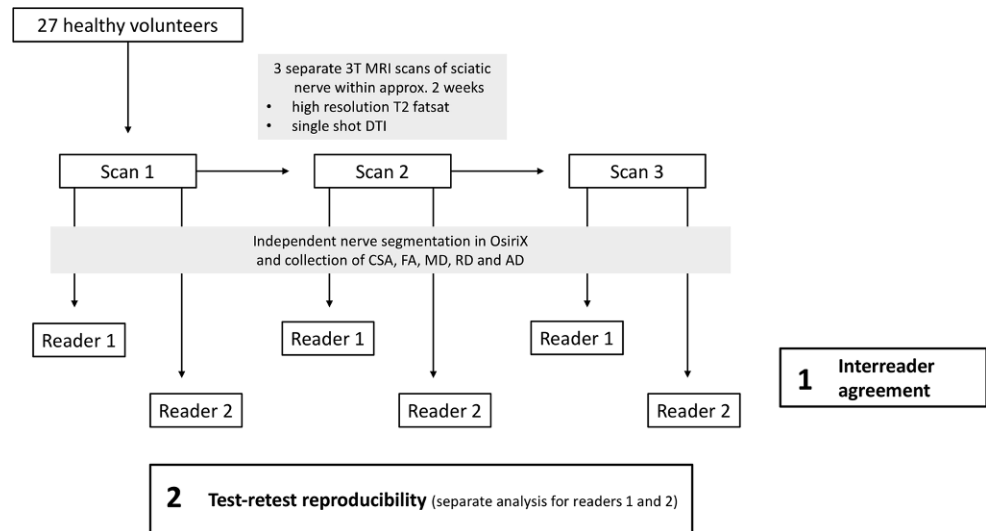
This study was approved by the institutional ethics committee. Written informed consent was obtained from all participants.

Study Subjects

For this study 28 healthy young men were recruited by public announcement and prospectively enrolled between October 2016 and August 2018. Of the participants one was excluded due to motion artifacts during the study so that 27 participants were finally analyzed. The mean age was 24.2 ± 3.2 years (range 19–30 years), mean weight was 76.6 ± 7.6 kg, mean height was 180 ± 5 cm, mean body mass index (BMI) was 23.8 ± 2.2 . The dominant side was right in 20/27 and left in 7/27.

All participants underwent three MR scans on different days, all of which were analyzed by two independent readers (flowchart in Fig. 1). Inclusion criteria were male sex and age between 18 and 35 years. Exclusion criteria were any known diseases of the peripheral nervous system as well as general contraindications for MRI and severe motion artifacts in the MRI examination.

Fig. 1 Flowchart of study design. A total number of 27 healthy volunteers underwent MR neurography on 3 separate days, each covering the sciatic nerve in the same anatomical region. Postprocessing and parameter readout was performed by two independent observers. Subsequently, statistical analyses were performed to calculate interreader agreement and test-retest reproducibility



MR Imaging

All participants were repeatedly examined at a 3.0T MR scanner (Magnetom Prisma-FIT, Siemens Healthcare, Erlangen, Germany) on 3 separate days. The mean timespan between the scanning sessions was 5.7 ± 2.3 days between scans 1 and 2 and 4.93 ± 1.11 days between scans 2 and 3. All participants underwent MR neurography of the dominant leg and were examined in a supine position with feet first and legs extended. A 15-channel transmit-receive knee coil (Siemens Healthineers, Erlangen, Germany) was placed at mid-thigh to distal thigh level and great care was taken that the distal end of the coil was approximately corresponding to the superior patella margin. This position not only allowed high reproducibility but also made it possible to perform high resolution imaging of the sciatic nerve on patients with larger thigh circumferences. In order to ensure a comfortable position and to avoid motion artifacts the thigh was immobilized with pads. In this position, the following image protocol was carried out:

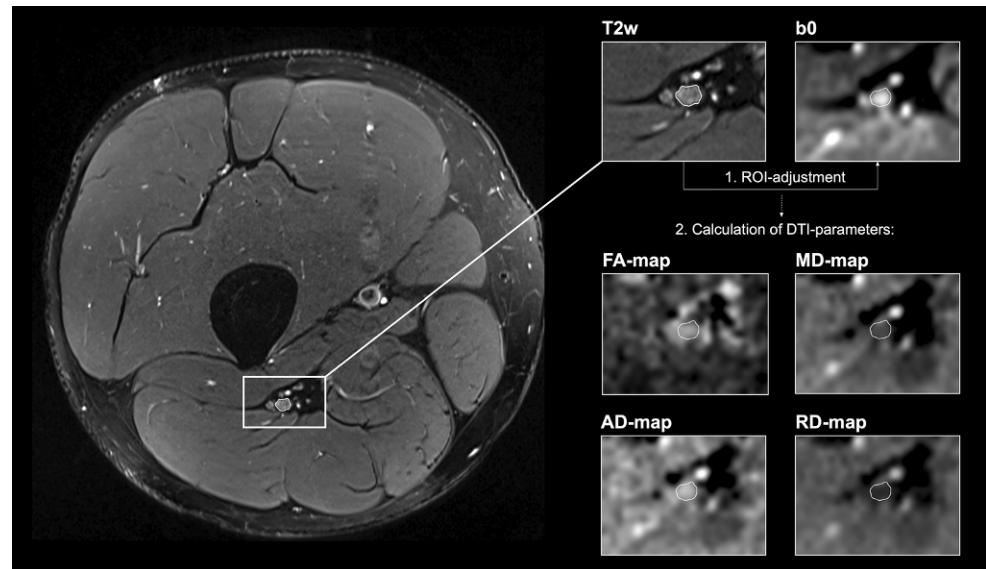
- i. Axial T2-weighted turbo spin echo (TSE) 2D sequence with spectral fat saturation, phase encoding in anteroposterior direction, TR/TE 8640/54 ms, matrix size 512×333 , phase resolution 65%, field of view $160 \times 160 \text{ mm}^2$, voxel size $0.3 \times 0.3 \times 3.5 \text{ mm}^3$, slice thickness 3.5 mm (distance factor 10%), number of slices 41, number of averages 3, acquisition time 6 min 48 s.
- ii. Axial single-shot spin-echo echo planar imaging (EPI) DTI with diffusion gradients in 20 directions (b -value = 0 and 800 s/mm^2) and fat suppression (spectrally adiabatic inversion recovery, SPAIR), phase encoding in anterior-posterior direction, TR/TE 7100/67 ms, matrix size 120×120 , field of view $160 \times 160 \text{ mm}^2$, voxel

size $1.3 \times 1.3 \times 3.5 \text{ mm}^3$, slice thickness 3.5 mm (distance factor 10%), number of slices 41, number of averages 2, acquisition time 5 min 42 s.

Image Postprocessing

Image analysis was performed by two independent readers (M.K. and F.P. with more than 5 and 3 years of experience in neuromuscular imaging, respectively) using the DICOM-viewer OsiriX (Pixmeo Sarl, Bernex, Switzerland). Image quality was visually assessed by both readers and one participant was excluded in consensus due to severe motion artifacts, and thus the data of 27 participants were analyzed in this study. All analyses in this study refer to these 27 participants. A total of seven slices of interest were defined at mid-thigh to distal-thigh level in each scan in identical anatomical positions by F.P. and then used for independent segmentation by both readers. Segmentation of the sciatic nerve was performed in the anatomical T2-weighted image by freehand regions of interests (ROI) where nerve size and shape could be well-delineated (Fig. 2). Analysis of the sciatic nerve was restricted to the tibial portion to avoid inclusion of fat and connective tissue. The ROIs were then copied to the corresponding b_0 -image ($b=0$) and slightly adjusted, mostly downsized, to avoid partial volume artifacts from surrounding fat tissue and to compensate possible distortion artifacts [19]. The FA, MD, AD, and RD were obtained using the OsiriX plugin DTI map with a preset noise threshold of 14. For further analysis, DTI parameters were averaged from all 7 slices.

Fig. 2 Representative images of nerve segmentation. Segmentation of the tibial portion of the sciatic nerve was conducted independently by each reader using a free-hand ROI tool in OsiriX. ROIs were transferred onto co-registered DTI maps and slightly adjusted in the b0-image before DTI parameters were obtained. *FA* fractional anisotropy, *MD* mean diffusivity, *RD* radial diffusivity, *AD* axial diffusivity, *T2w* T2-weighted image



Statistical Analysis

All statistical analyses were performed using SPSS (Version 24; SPSS Inc, Chicago, IL, USA) or the computing environment R (Version 3.4.4; R Development Core Team, R Foundation for Statistical Computing, Vienna, Austria). Values are shown as mean \pm standard deviation unless otherwise stated. Analyses of variance were conducted with the factors readers and scans (for each reader) to test for differences. Subjects were considered as random effects. Bonferroni correction was applied to correct for multiplicity. To assess interreader reliability a single measurement, absolute agreement, two-way random effects model, ICC (2,1) according to Shrout and Fleiss [31], was used to calculate ICCs with 95% confidence intervals (CI). The ICCs and its 95% CIs for the test-retest reliability were estimated based on a single measurement, absolute agreement, two-way mixed effects model as suggested by Shrout and Fleiss. Based on Koo and Li, ICC values between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.9 were considered as moderate, good, and excellent agreement, respectively [32]. Furthermore, a Bland-Altman analysis for repeated measurements was performed and illustrated in Bland-Altman plots [33]. To estimate the interreader and test-retest

variability the standard error of measurements (SEM) according to Popović and Thomas was calculated [25].

Results

Averaged over all measurements the mean \pm standard deviation FA was 0.507 ± 0.05 . MD was $1308.5 \pm 162.4 \times 10^{-6} \text{ mm}^2/\text{s}$, RD was $905.6 \pm 145.4 \times 10^{-6} \text{ mm}^2/\text{s}$ and AD was $2114.1 \pm 219.2 \times 10^{-6} \text{ mm}^2/\text{s}$. Testing with ANOVA revealed no significant differences between mean values of different scans and different readers for all DTI parameters. Detailed descriptive statistics of DTI parameters for all three scans and both readers are presented in Table 1 and visualized as scatter dot plots in Supplementary Fig. 1.

The ICCs for interreader and test-retest reliability are presented in Table 2. For interreader reliability they ranged from 0.81 to 0.92 implying good to excellent agreement according to Koo and Li [32]. The ICCs for test-retest reliability had only slightly lower values and ranged from 0.76 to 0.92, also implying good to almost perfect agreement.

The SEM for interreader reliability and test-retest reliability are presented in Table 3. The SEM for FA was 0.02 for interreader and test-retest agreement, the SEM

Table 1 Descriptive statistics of DTI measurements

	Scan 1		Scan 2		Scan 3	
	Reader 1	Reader 2	Reader 1	Reader 2	Reader 1	Reader 2
FA	0.51 ± 0.05	0.51 ± 0.05	0.50 ± 0.04	0.50 ± 0.05	0.51 ± 0.05	0.51 ± 0.05
MD ($10^{-6} \text{ mm}^2/\text{s}$)	1304 ± 165	1296 ± 168	1336 ± 158	1322 ± 167	1301 ± 160	1293 ± 166
RD ($10^{-6} \text{ mm}^2/\text{s}$)	905 ± 144	894 ± 148	930 ± 144	920 ± 152	896 ± 145	889 ± 149
AD ($10^{-6} \text{ mm}^2/\text{s}$)	2098 ± 231	2099 ± 226	2145 ± 210	2125 ± 224	2108 ± 214	2110 ± 226

All values represent mean \pm standard deviation

FA fractional anisotropy; MD mean diffusivity; RD radial diffusivity; AD axial diffusivity

Table 2 Intraclass correlation coefficients (ICCs) for inter-reader and test-retest reliability

	Interreader	Test-retest	
		Reader 1	Reader 2
<i>FA</i>	0.81 (0.63–0.91)	0.78 (0.63–0.88)	0.76 (0.60–0.87)
<i>MD</i>	0.92 (0.84–0.96)	0.92 (0.85–0.96)	0.91 (0.83–0.95)
<i>RD</i>	0.91 (0.82–0.96)	0.91 (0.83–0.96)	0.90 (0.81–0.95)
<i>AD</i>	0.92 (0.83–0.96)	0.89 (0.81–0.95)	0.91 (0.83–0.95)

All values represent intraclass correlation coefficients (ICCs) with 95% confidence intervals
FA fractional anisotropy; *MD* mean diffusivity; *RD* radial diffusivity; *AD* axial diffusivity

Table 3 Standard error of measurement (SEM) for interobserver and intraobserver variability

	Interreader	Test-retest
<i>FA</i>	0.02	0.02
<i>MD</i> (10^{-6} mm ² /s)	49.12	48.68
<i>RD</i> (10^{-6} mm ² /s)	46.23	45.86
<i>AD</i> (10^{-6} mm ² /s)	70.14	70.14

All values represent the SEM based on Popović and Thomas [25]
FA fractional anisotropy; *MD* mean diffusivity; *RD* radial diffusivity;
AD axial diffusivity

for MD, RD, and AD ranged between 46.2×10^{-6} mm²/s and 70.1×10^{-6} mm²/s for interreader reliability and between 45.9×10^{-6} mm²/s and 70.1×10^{-6} mm²/s for test-retest reliability. These SEM values can also be put in relation to the overall mean values, which resulted in a measurement error of approximately 3.9% for FA, 3.8% for MD, 5.1% for RD and 3.3% for AD for interreader reliability and 3.9% for FA, 3.7% for MD, 5.1% for RD and 3.3% for AD for test-retest reliability.

Bland-Altman plots for interreader reliability of all DTI parameters are shown in Fig. 3. Bland-Altman plots for test-retest reliability are shown in Fig. 4 (for FA) and in Supplementary Figs. 2–4 (for MD, RD and AD). As shown in the Bland-Altman plots, measurement error mainly occurs as a random fluctuation and not as a systematic error between scans or readers (bias). Moreover, visual assessment of Bland-Altman plots does not reveal a proportional bias (e.g. systematically higher measurement error for higher or lower FA, MD, or AD values).

Discussion

This study assessed the interreader and test-retest reliability of peripheral nerve DTI in a cohort of 27 healthy participants. The standard error of measurement (SEM) for all major DTI parameters is reported and thus orientation values are provided, which quantitatively describe the imprecision associated with repeated DTI measurements.

Until now, precision of peripheral nerve DTI has been assessed in several studies [15–22], examining the median nerve [15, 18, 19], the brachial plexus [16, 17], the tibial

and peroneal nerve [20], the sciatic nerve [21] as well as the lumbar and sacral nerves [22]. The design of these studies was heterogeneous and while most assessed interreader reliability, only three [16, 20, 22] assessed the test-retest reliability by repeated MR scans with identical parameters. All existing studies came to the conclusion that peripheral nerve DTI is principally reliable, and this conclusion is mainly based on the analysis of intraclass correlation coefficients (ICCs). The findings of the present study are in line with this conclusion as ICCs between 0.81 and 0.92 are reported for interreader and between 0.76 and 0.91 for test-retest reliability, which can be interpreted as good to excellent agreement [32].

The ICC is a dimensionless value between 0 and 1 and its calculation is based on the relation of the variance of the examined parameter in the population and its variance due to measurement error [25]. While ICC allows a general classification of reliability in terms of broad categories, it always has to be considered in conjunction with the variance of the assessed parameter in the examined population and not as an inherent property of the test itself.

The SEM is a statistical parameter that describes measurement precision independently of the variance in the population [24, 25]. It represents the standard deviation of the measurement error and while some SEM values have been briefly reported for interreader FA and apparent diffusion coefficient (ADC) values of the brachial plexus in one study [17], a systematic assessment of SEM in peripheral nerve DTI was still lacking. This study reports SEM values for both interreader and test-retest reliability of all major DTI parameters. The results of the present study may serve as orientation values when interpreting DTI parameters on an individual patient level.

Assuming a normal distribution of the measurement error, the SEM may be used to calculate CI around a measured value that cover the true value with 95% probability (95% CI: measured value $\pm 1.96 \times$ SEM) [25]. Moreover, SEM allows calculation of the minimum detectable difference (MDD = $2.8 \times$ SEM), which might be particularly useful when DTI parameters are used as biomarkers for follow-up of neuropathies. When the change of the measured parameter is larger than the MDD, it is likely due to a true change in the measured parameter and thus suggestive for

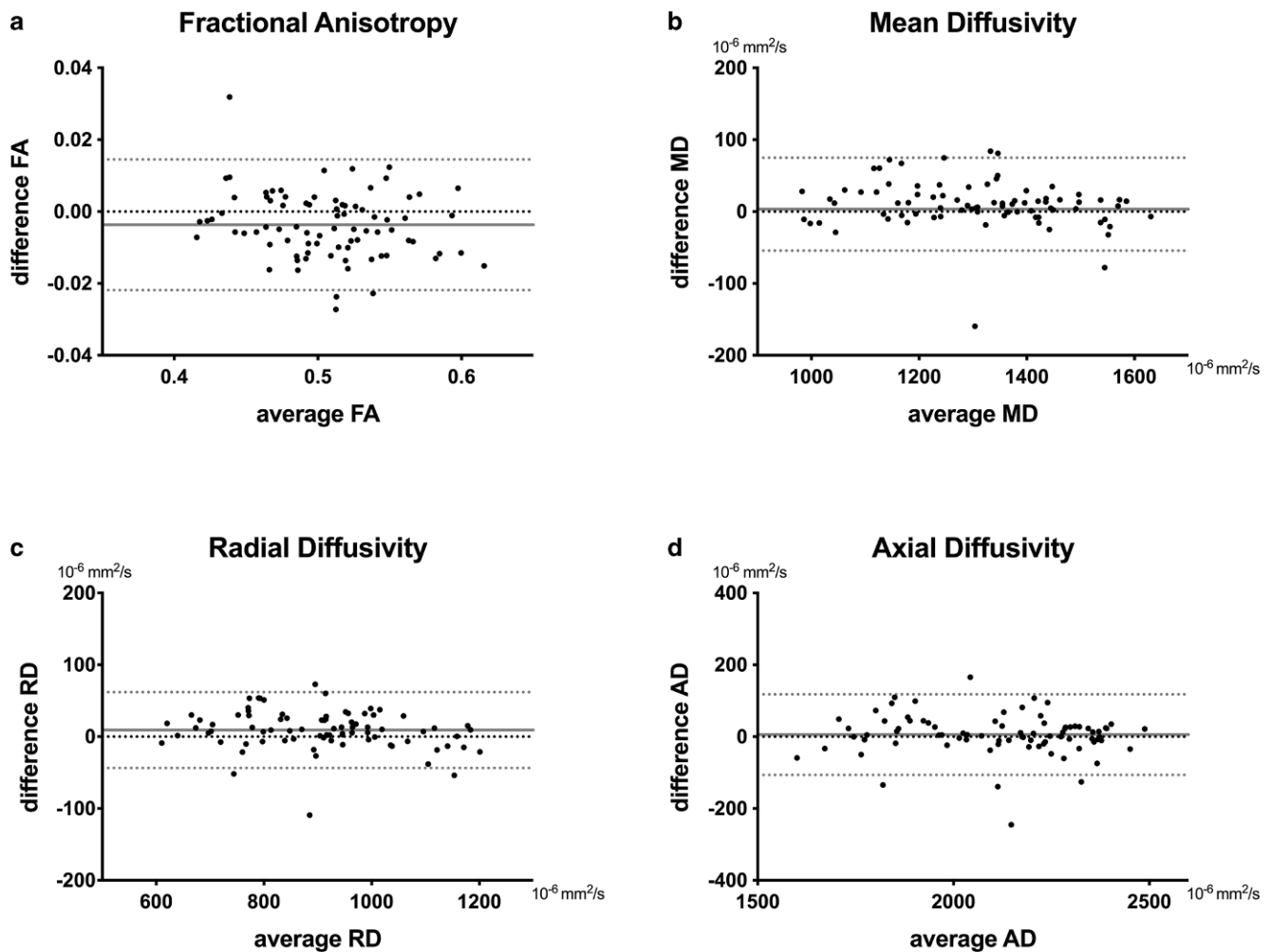


Fig. 3 Bland-Altman plots for assessment of interreader reliability of all DTI parameters with single values for bias (*grey line*) as well as for upper and lower limits of agreement (*dotted lines*). **a** FA fractional anisotropy, **b** MD mean diffusivity, **c** RD radial diffusivity, **d** AD axial diffusivity

a true nerve pathology. Otherwise the changes would probably only indicate a fluctuation in the observed parameter due to an accumulation of errors during recording (e.g. bias due to signal to noise ratio, partial volume effect that increases as nerve size decreases, motion artifacts) and/or postprocessing (e.g. segmentation errors) [25].

Among all DTI parameters, the FA is the most commonly reported parameter and regarded as a general measure of nerve tissue injury. In this study an SEM=0.02 was found for both test-retest and interreader reliability of FA. This implies an MDD of $2.8 \times 0.02 = 0.06$. These values may be regarded in conjunction with mean values of peripheral nerve FA that have been reported for various peripheral nerve diseases in multiple studies (see Table 4; [3–8, 10, 14, 30, 34–48]). In these studies, mean FA differences between healthy and diseased nerves ranged between 0.06 and 0.25, thus they are at least in the range of the MDD reported here.

The authors are aware of at least two longitudinal studies of peripheral nerve DTI in humans. Simon et al. assessed

DTI as a method of disease progression in amyotrophic lateral sclerosis (ALS) and found tibial nerve FA to be reduced from 0.434 (baseline) to 0.415 (after 6 months) [49]. While this difference (0.019) was significant on a group comparison level, a similar difference in a single patient, in a hypothetical examination setting like the present study, would be below the MDD and might therefore not be reliably differentiated from pure measurement error. Lichtenstein et al. on the other hand found that FA values in a clinically stable cohort of CIDP patients remained nearly unchanged over a follow-up period of 6 months and were still significantly lower compared to healthy controls (0.34 ± 0.01 at baseline; 0.35 ± 0.02 after 6 months; healthy controls 0.45 ± 0.01) [50]. To further elucidate the potential of DTI as a tool for individual patient follow-up, more longitudinal studies are needed that examine various pathologies and take into account the imprecision on an individual patient level.

This study has some limitations. First, the examined cohort was healthy and demographic variables were relatively

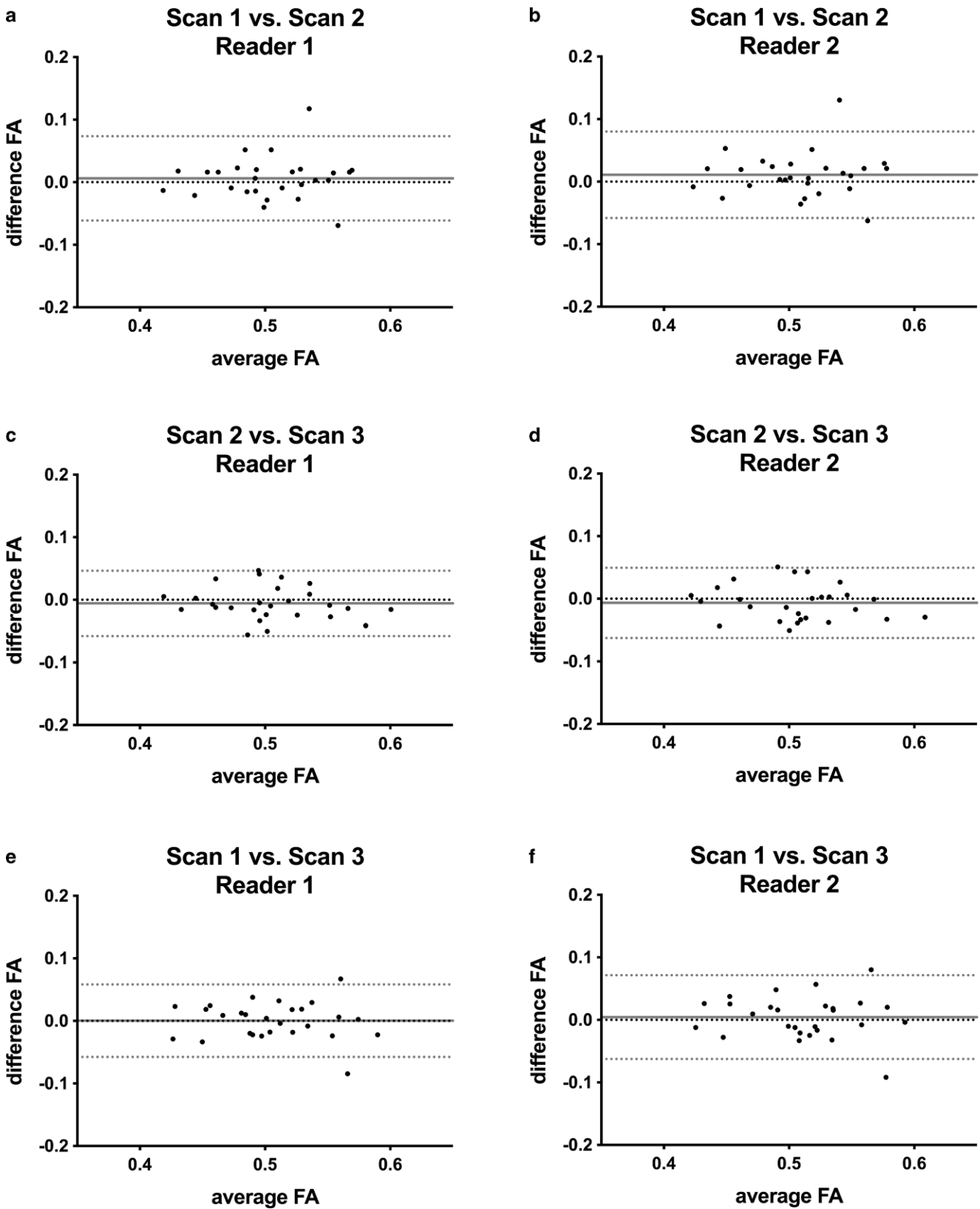


Fig. 4 Bland-Altman plots for assessment of test-retest reliability of fractional anisotropy (FA) values for both observers and all three scans, respectively. Plots as typically illustrated with single values for bias (*grey line*) as well as for upper and lower limits of agreement (*dotted lines*). **a, b** Scan 1 vs. (versus) Scan 2; **c, d** Scan 2 vs. Scan 3; **e, f** Scan 1 vs. Scan 3 for Reader 1 and 2, respectively

Table 4 Fractional anisotropy in various diseases as reported in previous studies of peripheral nerve DTI (nonexhaustive)

Authors	Examined pathology	Examined nerve	Healthy controls: (mean FA \pm SD)	Disease: (mean FA \pm SD)	Difference mean FA (healthy—disease)
Godel et al. 2019 [14]	Anterior interosseous nerve syndrome	Diseased fascicles of median nerve	0.58 \pm 0.08	0.33 \pm 0.08	0.25
Kim et al. 2019 [34]	Charcot-Marie-Tooth type 1A	Sciatic nerve (4 levels)	0.59 \pm 0.11	0.48 \pm 0.09	0.11
Schneider et al. 2019 [35]	Nonsystemic vasculitic neuropathy	Tibial nerve	0.42 \pm 0.01	0.32 \pm 0.02	0.10
Klauser et al. 2018 [36]	Carpal tunnel syndrome	Median nerve	0.62	0.54	0.08
Cao et al. 2018 [37]	Guillain-Barré syndrome	Tibial nerve	0.60 \pm 0.04	0.51 \pm 0.05	0.09
Wu et al. 2018 [38]	Chronic constriction injury	Sciatic nerve (rabbit)	0.57 \pm 0.04	0.46 \pm 0.04 (10-week follow-up)	0.11
Razek et al. 2017 [39]	Carpal tunnel syndrome	Median nerve	0.43 \pm 0.1	0.48 \pm 0.01	0.05
Vaeggemose et al. 2017 [30]	Charcot-Marie-Tooth type 1A	Sciatic nerve	0.48 ^a	0.35 ^a	0.13 ^a
Vaeggemose et al. 2017 [5]	Diabetic polyneuropathy (DPN)	Sciatic and tibial nerves	0.48 \pm 0.06 0.42 \pm 0.06	0.38 \pm 0.04 0.31 \pm 0.08 (severe DPN)	0.11–0.12
Haakma et al. 2017 [4]	Multifocal motor neuropathy (MMN) and amyotrophic lateral sclerosis (ALS)	Median and ulnar nerves	0.44 \pm 0.04	0.43 \pm 0.05 (ALS) 0.44 \pm 0.04 (MMN)	No difference
Wu et al. 2017 [40]	Diabetic polyneuropathy	Tibial nerve	0.59 \pm 0.64	0.53 \pm 0.52	0.06
Kronlage et al. 2017 [3]	Chronic inflammatory demyelinating polyneuropathy (CIDP)	Mean of multiple nerves	0.52 \pm 0.04	0.42 \pm 0.08	0.10
Bernabeu et al. 2016 [41]	Different types of sciatic neuropathy (e.g. entrapment, autoimmune)	Sciatic nerve	0.54 \pm 0.02	0.44 \pm 0.08	0.12
Markvardsen et al. 2016 [42]	CIDP	Sciatic nerve	0.48	0.42	0.06
Breitenseher et al. 2015 [10]	Ulnar nerve entrapment	Ulnar nerve	Significant reduction of ulnar FA values at the retrocondylar sulcus (no mean values are shown)		
Breckwoldt et al. 2015 [8]	Peripheral neuropathies of various etiologies	Various nerves	0.59 \pm 0.01	0.51 \pm 0.01	0.08
Jengojan et al. 2015 [7]	Acute radial nerve entrapment	Radial nerve	0.52 \pm 0.12	0.71 \pm 0.1	0.19
Shi et al. 2015 [43]	Nerve root compression	Lumbar nerve roots	0.22 \pm 0.02	0.20 \pm 0.04	0.02
Brienza et al. 2014 [44]	Carpal tunnel syndrome	Median nerve	0.59 \pm 0.01	0.36 \pm 0.06	0.23
Bäumer 2014 [45]	Subclinical ulnar neuropathy	Ulnar nerve	0.51 \pm 0.09	0.41 \pm 0.07	0.10
Mathys et al. 2013 [46]	Polyneuropathy of different origin (e.g. autoimmune)	Sciatic nerve	0.56	0.43	0.13
Chhabra et al. 2013 [47]	Peripheral nerve tumors and tumor-like conditions	Various nerves	0.34 \pm 0.1	0.22 \pm 0.1	0.12
Guggenberger et al. 2012 [6]	Carpal tunnel syndrome	Median nerve at various locations	0.55 \pm 0.01 (pisi-form bone)	0.44 \pm 0.01 (pisi-form bone)	0.11
Tanitame et al. 2012 [48]	Polyneuropathy with age-corrected FA	Tibial nerve	0.52	0.44	0.08

^a These values are given as median instead of mean values

Mean values for fractional anisotropy (FA) as found in the literature for various diseases are given for disease groups and healthy controls. Measures of precision as the standard error of measurement (SEM) and the minimum detectable difference (MDD) may be interpreted in conjunction with these values.

homogeneous. It was intended to assess measurement error under optimal conditions and it was decided to conduct this study in a young collective, since in our experience they are easier to position and exhibit less motion artifacts than many older subjects. A lower variance of the DTI parameters in the cohort than in a less homogeneous group may have led to lower ICC values. A larger study including older subjects and also participants with peripheral nerve disease would be desirable. While this cohort consisted of male participants the results are not expected to depend on the subjects' sex, as it has been shown that DTI parameters do not systematically differ between sexes [9]. Second, analysis was restricted to the sciatic nerve, which is well-suited for MRN. Analysis of other nerves with smaller caliber or oblique course, like in the upper extremities, may lead to higher measurement error. Therefore, the values given should be primarily used when performing MRN of the sciatic nerve. Moreover, patient positioning, MR examination and nerve segmentation were performed by experienced radiologists and the number of readers was two. Conditions in a clinical setting might differ and therefore the values of measurement error presented here should be considered as minimum values. Furthermore, the results are dependent on the segmentation method itself. It was chosen to perform the initial segmentation on high-resolution anatomical T2-weighted images where nerve morphology was accurately depicted and then the ROIs were transferred to the b0-image where they were adjusted. This method may use the advantages of high-resolution morphologic imaging while also compensating possible distortion artifacts. A direct segmentation based on the diffusion maps, e.g. FA map, would have the disadvantage that possible pathologies within the nerve could unintentionally be excluded due to their potentially reduced FA. Following segmentation, all DTI parameters were averaged from seven adjacent slices, similar to previous studies [3, 49, 50]. While averaging of multiple slices minimizes noise and segmentation errors, a relatively small coverage in the z-direction of approximately 2.7 cm reduces possible effects from B1-field inhomogeneities and from potential physiological alterations of DTI parameters along the course of the nerve [34, 51]. Also, despite being relatively small the coverage is large enough for a robust re-identification of a nerve segment in relation to anatomical landmarks in follow-up examinations. The segmentation technique could be similarly applied in a clinical setting, e.g. in the follow-up of systemic neuropathies. In the future autosegmentation methods will probably become more important since they have already been applied to the sciatic nerve with good accuracy [52]. The full potential of these methods remains to be elucidated. Finally, this study represents a single vendor study. Assessments of other manufacturers, studies with other sequence parameters, subject populations, segmenta-

tion methods and postprocessing software are desirable and the values of measurement accuracy presented here should be interpreted with caution.

In summary, this study assessed interreader and test-retest reliability of peripheral nerve DTI in a cohort of 27 healthy volunteers. This study confirms adequate reliability of peripheral nerve DTI in principle and by reporting SEM for all major DTI parameters it also quantitatively assesses the imprecision that is associated with different readers or multiple scans. The values presented here may serve as first orientation values of measurement precision when interpreting parameters of sciatic nerve DTI.

Funding Supported in part by the German Research Council (SFB 1118, S.H.; SFB 1158, M.B.).

Compliance with ethical guidelines

Conflict of interest F. Preisner, P. Bäumer, M. Wehrstein, B. Friedmann-Bette, M. Hackbusch, S. Heiland, M. Bendszus and M. Kronlage declare that they have no competing interests.

Ethical standards All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1975 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

- Hagmann P, Jonasson L, Maeder P, Thiran JP, Wedeen VJ, Meuli R. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics*. 2006;26(Suppl 1):S205–23.
- Mori S, Zhang J. Principles of diffusion tensor imaging and its applications to basic neuroscience research. *Neuron*. 2006;51:527–39.
- Kronlage M, Pitarokoili K, Schwarz D, Godel T, Heiland S, Yoon MS, Bendszus M, Bäumer P. Diffusion tensor imaging in chronic inflammatory demyelinating polyneuropathy: diagnostic accuracy and correlation with electrophysiology. *Invest Radiol*. 2017;52:701–7.
- Haakma W, Jongbloed BA, Froeling M, Goedee HS, Bos C, Leemans A, van den Berg LH, Hendrikse J, van der Pol WL. MRI shows thickening and altered diffusion in the median and ulnar nerves in multifocal motor neuropathy. *Eur Radiol*. 2017;27:2216–24.
- Vaeggemose M, Pham M, Ringgaard S, Tankisi H, Ejkskjær N, Heiland S, Poulsen PL, Andersen H. Magnetic resonance neurography visualizes abnormalities in sciatic and tibial nerves in patients with type 1 diabetes and neuropathy. *Diabetes*. 2017;66:1779–88. <https://doi.org/10.2337/db16-1049>
- Guggenberger R, Markovic D, Eppenberger P, Chhabra A, Schiller A, Nanz D, Prüssmann K, Andreisek G. Assessment of median nerve with MR neurography by using diffusion-tensor imaging: normative and pathologic diffusion values. *Radiology*. 2012;265:194–203.
- Jengojan S, Kovar F, Breitensteher J, Weber M, Prayer D, Kasprian G. Acute radial nerve entrapment at the spiral groove: detection by DTI-based neurography. *Eur Radiol*. 2015;25:1678–83.

8. Breckwoldt MO, Stock C, Xia A, Heckel A, Bendszus M, Pham M, Heiland S, Bäumer P. Diffusion tensor imaging adds diagnostic accuracy in magnetic resonance neurography. *Invest Radiol.* 2015;50:498–504.
9. Kronlage M, Schwehr V, Schwarz D, Godel T, Uhlmann L, Heiland S, Bendszus M, Bäumer P. Peripheral nerve diffusion tensor imaging (DTI): normal values and demographic determinants in a cohort of 60 healthy individuals. *Eur Radiol.* 2018;28:1801–8.
10. Breitensteher JB1, Kranz G, Hold A, Berzaczy D, Nemeš SF, Sycha T, Weber M, Prayer D, Kasprian G. MR neurography of ulnar nerve entrapment at the cubital tunnel: a diffusion tensor imaging study. *Eur Radiol.* 2015;25:1911–8.
11. Hiltunen J, Kirveskari E, Numminen J, Lindfors N, Göransson H, Hari R. Pre- and post-operative diffusion tensor imaging of the median nerve in carpal tunnel syndrome. *Eur Radiol.* 2012;22:1310–9.
12. O'Donnell LJ, Westin CF. An introduction to diffusion tensor image analysis. *Neurosurg Clin N Am.* 2011;22:185–96.
13. Heckel A, Weiler M, Xia A, Ruetters M, Pham M, Bendszus M, Heiland S, Baeumer P. Peripheral nerve diffusion tensor imaging: assessment of axon and myelin sheath integrity. *PLoS One.* 2015;10:e130833.
14. Godel T, Pham M, Kele H, Kronlage M, Schwarz D, Brunée M, Heiland S, Bendszus M, Bäumer P. Diffusion tensor imaging in anterior interosseous nerve syndrome—functional MR neurography on a fascicular level. *Neuroimage Clin.* 2019;21:101659.
15. Andreisek G, White LM, Kassner A, Sussman MS. Evaluation of diffusion tensor imaging and fiber tractography of the median nerve: preliminary results on intrasubject variability and precision of measurements. *AJR Am J Roentgenol.* 2010;194:W65–72.
16. Ho MJ, Manoliu A, Kuhn FP, Stieltjes B, Klarhöfer M, Feiweier T, Marcon M, Andreisek G. Evaluation of reproducibility of diffusion tensor imaging in the brachial plexus at 3.0T. *Invest Radiol.* 2017;52:482–7.
17. Tagliafico A, Calabrese M, Puntoni M, Pace D, Baio G, Neumaier CE, Martinoli C. Brachial plexus MR imaging: accuracy and reproducibility of DTI-derived measurements and fibre tractography at 3.0-T. *Eur Radiol.* 2011;21:1764–71.
18. Guggenberger R, Nanz D, Bussmann L, Chhabra A, Fischer MA, Hodler J, Pfirrmann CW, Andreisek G. Diffusion tensor imaging of the median nerve at 3.0T using different MR scanners: agreement of FA and ADC measurements. *Eur J Radiol.* 2013;82:e590–6.
19. Guggenberger R, Nanz D, Puipe G, Rufibach K, White LM, Sussman MS, Andreisek G. Diffusion tensor imaging of the median nerve: intra-, inter-reader agreement, and agreement between two software packages. *Skeletal Radiol.* 2012;41:971–80.
20. Simon NG, Lagopoulos J, Gallagher T, Kliot M, Kiernan MC. Peripheral nerve diffusion tensor imaging is reliable and reproducible. *J Magn Reson Imaging.* 2016;43:962–9.
21. Keller S, Wang ZJ, Golsari A, Kim AC, Kooijman H, Adam G, Yamamura J. Feasibility of peripheral nerve MR neurography using diffusion tensor imaging adapted to skeletal muscle disease. *Acta Radiol.* 2018;59:560–8.
22. Haakma W, Hendrikse J, Uhrenholt L, Leemans A, Warner Thorup Boel L, Pedersen M, Froeling M. Multicenter reproducibility study of diffusion MRI and fiber tractography of the lumbosacral nerves. *J Magn Reson Imaging.* 2018;48:951–63.
23. Lee KM, Lee J, Chung CY, Ahn S, Sung KH, Kim TW, Lee HJ, Park MS. Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. *Clin Orthop Surg.* 2012;4:149–55.
24. Stratford PW, Goldsmith CH. Use of the standard error as a reliability index of interest: an applied example using elbow flexor strength data. *Phys Ther.* 1997;77:745–50.
25. Popović ZB, Thomas JD. Assessing observer variability: a user's guide. *Cardiovasc Diagn Ther.* 2017;7:317–24.
26. Godel T, Bäumer P, Farschtschi S, Gugel I, Kronlage M, Hofstadler B, Heiland S, Gelderblom M, Bendszus M, Mautner VF. Peripheral nervous system alterations in infant and adult neurofibromatosis type 2. *Neurology.* 2019;93:e590–8.
27. Jende JME, Hauck GH, Diem R, Weiler M, Heiland S, Wildemann B, Korporal-Kuhnke M, Wick W, Hayes JM, Pfaff J, Pham M, Bendszus M, Kollmer J. Peripheral nerve involvement in multiple sclerosis: demonstration by magnetic resonance neurography. *Ann Neurol.* 2017;82:676–85.
28. Jende JME, Groener JB, Oikonomou D, Heiland S, Kopf S, Pham M, Nawroth P, Bendszus M, Kurz FT. Diabetic neuropathy differs between type 1 and type 2 diabetes: Insights from magnetic resonance neurography. *Ann Neurol.* 2018;83:588–98.
29. Jende JME, Groener JB, Rother C, Kender Z, Hahn A, Hilgenfeld T, Juerchott A, Preisner F, Heiland S, Kopf S, Pham M, Nawroth P, Bendszus M, Kurz FT. Association of serum cholesterol levels with peripheral nerve damage in patients with type 2 diabetes. *JAMA Netw Open.* 2019;2:e194798.
30. Vaeggemose M, Vaeth S, Pham M, Ringgaard S, Jensen UB, Tankisi H, Ejksjaer N, Heiland S, Andersen H. Magnetic resonance neurography and diffusion tensor imaging of the peripheral nerves in patients with Charcot-Marie-Tooth Type 1A. *Muscle Nerve.* 2017;56:E78–84.
31. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420–8.
32. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15:155–63.
33. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8:135–60.
34. Kim HS, Yoon YC, Choi BO, Jin W, Cha JG, Kim JH. Diffusion tensor imaging of the sciatic nerve in Charcot-Marie-tooth disease type I patients: a prospective case-control study. *Eur Radiol.* 2019;29:3241–52.
35. Schneider C, Sprenger A, Weiss K, Slebocki K, Maintz D, Fink GR, Henning TD, Lehmann HC, Lichtenstein T. MRI detects peripheral nerve and adjacent muscle pathology in non-systemic vasculitic neuropathy (NSVN). *J Neurol.* 2019;266:975–81.
36. Klauser AS, Abd Ellah M, Kremser C, Taljanovic M, Schmidle G, Gabl M, Cartes-Zumelzu F, Steiger R, Gizewski ER. Carpal tunnel syndrome assessment with diffusion tensor imaging: value of fractional anisotropy and apparent diffusion coefficient. *Eur Radiol.* 2018;28:1111–7.
37. Cao J, He B, Wang S, Zhou Z, Gao F, Xiao L, Luo X, Wu C, Gong T, Chen W, Wang G. Diffusion tensor imaging of tibial and common peroneal nerves in patients with Guillain-Barre syndrome: a feasibility study. *J Magn Reson Imaging.* 2019;49:1356–64.
38. Wu W, Niu Y, Kong X, Liu D, Long X, Shu S, Su X, Wang B, Liu X, Ma Y, Wang L. Application of diffusion tensor imaging in quantitatively monitoring chronic constriction injury of rabbit sciatic nerves: correlation with histological and functional changes. *Br J Radiol.* 2018;91:20170414.
39. Razeq AAKA, Shabana AAE, El Saied TO, Alrefey N. Diffusion tensor imaging of mild-moderate carpal tunnel syndrome: correlation with nerve conduction study and clinical tests. *Clin Rheumatol.* 2017;36:2319–24.
40. Wu C, Wang G, Zhao Y, Hao W, Zhao L, Zhang X, Cao J, Wang S, Chen W, Chan Q, Zhao B, Chhabra A. Assessment of tibial and common peroneal nerves in diabetic peripheral neuropathy by diffusion tensor imaging: a case control study. *Eur Radiol.* 2017;27:3523–31.
41. Bernabeu Á, López-Celada S, Alfaro A, Mas JJ, Sánchez-González J. Is diffusion tensor imaging useful in the assessment of the sciatic nerve and its pathologies? Our clinical experience. *Br J Radiol.* 2016;89:20150728.

42. Markvardsen LH, Vaeggemose M, Ringgaard S, Andersen H. Diffusion tensor imaging can be used to detect lesions in peripheral nerves in patients with chronic inflammatory demyelinating polyneuropathy treated with subcutaneous immunoglobulin. *Neuroradiology*. 2016;58:745–52.
43. Shi Y, Zong M, Xu X, Zou Y, Feng Y, Liu W, Wang C, Wang D. Diffusion tensor imaging with quantitative evaluation and fiber tractography of lumbar nerve roots in sciatica. *Eur J Radiol*. 2015;84:690–5.
44. Brienza M, Pujia F, Colaiacomo MC, Anastasio MG, Pierelli F, Di Biasi C, Andreoli C, Gualdi G, Valente GO. 3T diffusion tensor imaging and electroneurography of peripheral nerve: a morphofunctional analysis in carpal tunnel syndrome. *J Neuroradiol*. 2014;41:124–30.
45. Bäumer P, Pham M, Ruetters M, Heiland S, Heckel A, Radbruch A, et al. Peripheral neuropathy: detection with diffusion-tensor imaging. *Radiology*. 2014;273(1):185–93. <https://doi.org/10.1148/radiol.14132837>.
46. Mathys C, Aissa J, Meyer Zu Hörste G, Reichelt DC, Antoch G, Turowski B, Hartung HP, Sheikh KA, Lehmann HC. Peripheral neuropathy: assessment of proximal nerve integrity by diffusion tensor imaging. *Muscle Nerve*. 2013;48:889–96.
47. Chhabra A, Thakkar RS, Andreisek G, Chalian M, Belzberg AJ, Blakeley J, Hoke A, Thawait GK, Eng J, Carrino JA. Anatomic MR imaging and functional diffusion tensor imaging of peripheral nerve tumors and tumorlike conditions. *AJNR Am J Neuroradiol*. 2013;34:802–7.
48. Tanitame K, Iwakado Y, Akiyama Y, Ueno H, Ochi K, Otani K, Takasu M, Date S, Awai K. Effect of age on the fractional anisotropy (FA) value of peripheral nerves and clinical significance of the age-corrected FA value for evaluating polyneuropathies. *Neuroradiology*. 2012;54:815–21.
49. Simon NG, Lagopoulos J, Paling S, Pfluger C, Park SB, Howells J, Gallagher T, Kliot M, Henderson RD, Vucic S, Kiernan MC. Peripheral nerve diffusion tensor imaging as a measure of disease progression in ALS. *J Neurol*. 2017;264:882–90.
50. Lichtenstein T, Sprenger A, Weiss K, Slebocki K, Cervantes B, Karampinos D, Maintz D, Fink GR, Henning TD, Lehmann HC. MRI biomarkers of proximal nerve injury in CIDP. *Ann Clin Transl Neurol*. 2018;5:19–28.
51. Wako Y, Nakamura J, Eguchi Y, Hagiwara S, Miura M, Kawarai Y, Sugano M, Nawata K, Yoshino K, Toguchi Y, Masuda Y, Matsumoto K, Suzuki T, Orita S, Ohtori S. Diffusion tensor imaging and tractography of the sciatic and femoral nerves in healthy volunteers at 3T. *J Orthop Surg Res*. 2017;12:184.
52. Balsiger F, Steindel C, Arn M, Wagner B, Grunder L, El-Koussy M, Valenzuela W, Reyes M, Scheidegger O. Segmentation of peripheral nerves from magnetic resonance neurography: a fully-automatic, deep learning-based approach. *Front Neurol*. 2018;9:777.