Improvement of Peripheral Nerve Visualization Using a Deep Learning-based MR Reconstruction Algorithm

Kelly C. Zochowski
Hospital for Special Surgery
Ek Tsoon Tan
Hospital for Special Surgery
Erin C. Argentieri
Hospital for Special Surgery
Bin Lin
Hospital for Special Surgery
Alissa J. Burge
Hospital for Special Surgery
Sophie C. Queler
Hospital for Special Surgery
R. Marc Lebel
General Electric (United States)
Darryl B. Sneag (sneagd@hss.edu)
Hospital for Special Surgery

Research Article

Keywords: peripheral nerves, magnetic resonance imaging, deep learning, humans, artificial intelligence

Posted Date: May 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-493137/v1

Version of Record: A version of this preprint was published at Magnetic Resonance Imaging on January 1st, 2022. See the published version at https://doi.org/10.1016/j.mri.2021.10.038.

Abstract

Objective: To assess a new deep learning-based MR reconstruction method, “DLRecon,” for clinical evaluation of peripheral nerves.

Methods: Sixty peripheral nerves were prospectively evaluated in 29 patients (mean age: 49±16 years, 17 female) undergoing standard-of-care (SOC) MR neurography for clinically suspected neuropathy. SOC-MRIs and DLRecon-MRIs were obtained through conventional and DLRecon reconstruction methods, respectively. Two radiologists randomly evaluated blinded images for outer epineurium conspicuity, fascicular architecture visualization, pulsation artifact, ghosting artifact, and bulk motion.

Results: DLRecon-MRIs were likely to score better than SOC-MRIs for outer epineurium conspicuity (OR=1.9, p=0.007) and visualization of fasicular architecture (OR=1.8, p<0.001) and were likely to score worse for ghosting (OR=2.8, p=0.004) and pulsation artifacts (OR=1.6, p=0.004). There was substantial to almost-perfect inter-reconstruction method agreement (AC=0.73-1.00) and fair to almost-perfect interrater agreement (AC=0.34-0.86) for all features evaluated. DLRecon-MRI had improved interrater agreement for outer epineurium conspicuity (AC=0.71, substantial agreement) compared to SOC-MRIs (AC=0.34, fair agreement). In >80% of images, the radiologist correctly identified an image as SOC- or DLRecon-MRI.

Discussion: Outer epineurium and fasicular architecture conspicuity, two key morphological features critical to evaluating a nerve injury, were improved in DLRecon-MRIs compared to SOC-MRIs. Although pulsation and ghosting artifacts increased in DLRecon images, image interpretation was unaffected.

Introduction

MR neurography is challenging due to the small size of peripheral nerves, some less than 1-2 mm in maximal caliber,[1] and the need to evaluate both the outer epineurium and inner fascicular architecture.[2] High spatial resolution (<0.5mm), cross-sectional acquisition is therefore important. In the extremities, ~0.3 mm in-plane resolution is currently achieved at 3.0 Tesla field strength.[3] However, realizing this resolution within clinically reasonable scan times (<6 minutes) and with adequate signal-to-noise ratio (SNR) is challenging.[4]

Acceleration techniques to reduce scan time or increase spatial resolution include parallel imaging [5,6] and compressed sensing,[7] and are facilitated via high channel surface coils to improve SNR. However, acceleration methods incur SNR penalties due to under-sampling and noise amplification,[6] which may obscure relevant image details when high acceleration rates (beyond ~2x) are used. A high nerve-to-muscle contrast-to-noise ratio is also required, as nerves often course adjacent to or within muscles with similar contrast.

Another approach to obtaining diagnostic quality images within reasonable scans times is to leverage artificial intelligence (AI) for denoising,[8] super-resolution,[9] artifact reduction, and/or reconstruction of under-sampled data.[10] Broadly, AI-enhanced image reconstruction algorithms can operate alongside

conventional algorithms, including parallel imaging or partial Fourier, to compensate for noise amplification or blurring. Alternatively, AI has been used to reconstruct images directly from fully or undersampled k-space data. One such AI method, deep learning based AIR Recon DL (DLRecon),[11] operates on raw image data alongside conventional parallel imaging algorithms during image reconstruction. DLRecon is specifically designed to perform both image denoising and Gibbs ringing removal, ultimately producing images with high SNR and sharp edges.[11] The DLRecon AI was previously trained on a curated database of over 4 million training iterations and applies to two-dimensional (2D) imaging of any anatomy, contrast weighting, or coil configuration.[11] This technique was recently applied to thin-slice pituitary gland MRI[12] and late gadolinium enhancement in myocardial scar quanitification,[13] where DLRecon demonstrated similar or better diagnostic performance relative to the conventional MR image reconstruction method.

Given limitations of conventional image reconstruction methods for MR neurography and previously reported concerns about compromise of image fidelity with AI,[14] it was desirable to evaluate DLRecon's ability to improve image quality and its agreement with the existing, standard of care (SOC) image reconstruction method. We hypothesized the following: (i) DLRecon would enhance clinically relevant imaging features of peripheral nerves without increasing artifact; and (ii) DLRecon would improve inter-rater agreement of these measures compared to SOC-reconstruction.

Methods

Approval for Research in Human Subjects

This study was approved by the institutional review board of Hospital for Special Surgery and conducted in accordance with the Health Insurance Portability and Accountability Act. Informed consent was obtained from all individual participants.

Study Design and Study Population

In total, 29 subjects were prospectively enrolled and in these subjects 60 peripheral nerves were evaluated. These nerves were chosen for evaluation as they are commonly evaluated in clinical practice and were the largest in diameter within the imaged field-of-view. Written informed consent was obtained from all patients prior to imaging performed between February 2019 and November 2019.

Inclusion Criteria

All patients who presented to our institution for standard-of-care MR neurography evaluation for clinically suspected neuropathy were considered for study inclusion.

Exclusion Criteria

Exclusion criteria were standard MRI safety contraindications. No patients were excluded from the study.

Image Acquisition

All scans were performed on a 3.0 Tesla clinical scanner (MR750, GE Healthcare) using a 16-channel flexible coil (Neocoil). Axial, 2D intermediate-weighted fast spin echo sequences (FSE) were obtained as part of the institution's standard MR neurography protocol, using the following acquisition parameters optimized for each body part: echo time (TE): 21-39 ms, repetition time (TR): 3177-6522 ms, receiver bandwidth (RBW): 195.312-488.281 Hz/pixel, field of view: 80-320 mm, acquisition matrix: (256 or 512)x(256-512), slice thickness: 2.0-4.5 mm, echo train length (ETL): 8-21, number of excitations (NEX): 1-4, parallel imaging factors between 1.5 and 2. No compressed-sensing acquisition was utilized.

Image Reconstruction

In addition to the untouched SOC reconstruction (i.e. images immediately derived from the scanner), the raw data was retrospectively reconstructed with a deep convolution neural network, i.e. "DLRecon", [11] a vendor-provided software installed on the scanner. This neural network accepts raw unfiltered complex valued image inputs and outputs images with higher SNR and reduced truncation artifacts by utilizing a feed-forward approach.[11] Gibbs ringing that occurs near sharp edges is also removed by the neural network, resulting in increased image sharpness.[11]

DLRecon was previously trained using pairs of images containing conventional MR images and 'near-perfect' images, defined as those with high resolution, minimal ringing, and very low noise levels.[11] Four million unique image pairs were employed in this supervised learning approach, and image augmentations such as rotations, flips, intensity gradients, phase manipulations, and Gaussian noise were used to increase the robustness of the training set.[11] The training images were diverse, allowing generalizabitlity of DLRecon's application across anatomical locations.[11] Additionally, the network was trained using a gradient backpropogation and ADAM optimizer.[11,15]

Image Analysis

Anonymized images were evaluated on a picture archiving and communication system (PACS) (Sectra V18.1, Sectra AB) by two board-certified radiologists: reader 1 (DBS) with 6 years of dedicated MR neurography experience, and reader 2 (AJB) with 10 years of general musculoskeletal MRI experience. Both readers underwent a training session to establish grading consensus by reviewing both SOC- and DLRecon-MRI images from 10 separate datasets not included in the analysis. Study images were randomized prior to evaluation by the 2 readers, who remained blinded with respect to the post-processing method. Each reader independently scored images for outer epineurium conspicuity and visualization of fascicular architecture by reviewing the entire volume for slices that best demonstrated the nerve in an orthogonal plane, as this is the best plane to visualize fascicular architecture. For evaluation of pulsation artifact, ghosting artifact, and bulk motion, readers considered all slices. Outer epineurium conspicuity was defined by the number of distinct borders visualized between the nerve and immediately surrounding perineural fat (maximum 4: anterior, posterior, medial, and lateral). Fascicular architecture was graded using the following Likert scale: 1-poor visualization, 2-average visualization, 3-

good visualization, 4-excellent visualization. The presence of pulsation artifact, ghosting artifact, and bulk motion was graded using the following scale: 0-none, 1-mild, 2-moderate, 3-severe. Additionally, each radiologist was asked to 'guess' as to whether each image dataset was processed with SOC-MRI or DLRecon-MRI to determine the extent of potential bias from perceiving image texture differences.

Statistical Analysis

Statistical analyses were performed by a biostatistician (BL) with 5 years of experience. Odds ratios (OR) and 95% confidence intervals (CI) obtained from marginal ordinal logistic regression models estimated with generalized estimating equation were used to evaluate for differences in grades between DLRecon- and SOC- MRIs. Patients were treated as the repeated factor to account for any within-patient correlations between image type as well as patients from whom more than one nerve was evaluated in their exam. Given that ORs were calculated as a comparison of DLRecon-MR and SOC-MR images, an OR of 1 was interpreted as no difference between DLRecon- and SOC-MRI; an OR >1 was interpreted as the DLRecon-MRI being more likely to have a higher grade than the SOC-MRI; and an OR <1 was interpreted as the SOC-MRI being more likely to have a higher grade than the DLRecon-MRI. Statistical significance was set *a priori* to $p< 0.05$.

Agreement between DLRecon- and SOC-MRI grades for each reader (inter-reconstruction agreement) was analyzed using ordinal-weighted Gwet's agreement coefficients (AC). Clustered bootstrap confidence intervals were used to account for patients who had more than 1 nerve examined. Strength of agreement was determined using the following scale: <0 = poor, 0.00-0.2 = slight, 0.21-0.40 = fair, 0.41-0.60 = moderate, 0.61-0.80 = substantial, and 0.81-1 = almost-perfect.[16] The interrater agreement was analyzed in the same manner. Statistical analyses were performed with SAS v. 9.4 (SAS Institute).

Results

The study cohort included 12 male (mean age, 49.5 ± 17.8 years old) and 17 female (mean age, 49.4 ± 15.2 years old) patients with no statistical age difference between sexes (p=0.989) (Table 1). A total of 60 axial 2D intermediate-weighted FSE sequences were evaluated in the 29 subjects, with 18 subjects undergoing evaluation for more than one nerve or anatomic location. Overall, 13 thighs, 4 knees, 3 lower legs, 3 elbows, 2 forearms, 1 arm, 1 wrist, 1 hand, and 1 ankle were scanned. Within these scans, 21 (35%) sciatic, 16 (27%) tibial, 16 (27%) median, and 7 (12%) ulnar nerves were graded.

DLRecon-MRI Versus SOC-MRI

The distribution of image quality features evaluated on SOC and DLRecon MRIs for reader 1 and reader 2 is demonstrated in Table 2. When directly comparing the rating of image quality features in DLRecon and SOC using reader 1's grades (Table 3), DLRecon was more likely to be scored higher for fascicular architecture (OR=1.8, p=0.057) and outer epineurium conspicuity (OR=1.4, p=0.478) but worse for pulsation artifact (OR=1.2, p=0.141) and bulk motion (OR=1.2, p=0.719), although these results did not achieve statistical significance. One result for reader 1 that did demonstrate statistical significance was

increased severity of ghosting artifact (OR=1.7, p=0.001, Figure 1). Reader 2 however, scored DLRecon significantly higher for both outer epineurium (OR=1.9, p=0.007, Figure 2) and fascicular architecture (OR=1.8, p<0.001, Figures 3 and 4), and significantly worse for pulsation (OR=1.6, p=0.004) and ghosting artifact (OR=2.8, p<0.001) (Figure 1).

Inter-reconstruction and Interrater Agreement

For inter-reconstruction agreement, each reader had substantial to almost-perfect agreement for all imaging features and artifacts evaluated (AC=0.73-1.00) (Table 4). However, interrater agreement was lower and variable. There was substantial interrater agreement for visualization of fascicular architecture in both SOC and DLRecon (AC=0.73-0.75) and almost-perfect agreement for bulk motion in both SOC and DLRecon (AC=0.84-0.86). However, for outer epineurium conspicuity there was substantial interrater agreement in DLRecon (AC=0.71) but only fair agreement in SOC (AC=0.34). Interrater agreement for pulsation artifact was substantial to almost-perfect (AC=0.76-0.86), and moderate to substantial for ghosting artifact (AC=0.57-0.64).

Reader Impression of Image as SOC-MRI or DLRecon-MRI

Despite blinding, readers 1 and 2 correctly identified images as either SOC or DLRecon 83% and 91% of the time, respectively.

Discussion

This study demonstrated efficacy of an AI-reconstruction algorithm (DLRecon) to improve peripheral nerve evaluation on MRI. Specifically, the integrity of two morphologic features (the outer epineurium and fascicular architecture) that are critical to determine the presence and extent of peripheral nerve injury, were shown to be more conspicuous with DLRecon relative to SOC reconstruction. Study findings were concordant with the expected outcomes of the algorithm, namely denoising and improved sharpness to enhance edge definition.

One potential concern of AI-based reconstruction is reduced image fidelity, resulting in over-smoothing of the images and loss of image details. This was partially addressed by evaluating both improvement (based on the OR) and inter-reconstruction agreement to determine whether there was significant bias from either the training dataset or algorithm. Conspicuity of nerve features was improved with DLRecon and yet inter-reconstruction agreement was high, suggesting improved image quality with no substantial change in interpretation. Apparent changes in image smoothness with DLRecon could bias reader interpretation. In fact, despite blinding, readers were mostly able to correctly identify data sets as either SOC- or DLRecon-MRI. This suggests the presence of noticeable image texture differences between the two methods; anecdotally, readers observed visible noise reduction and image sharpening in DLRecon compared to SOC images in the muscle and bone. Nonetheless, any potential bias was qualified by the substantial agreement between the two reconstruction methods.

DLRecon increased pulsation artifacts and ghosting artifacts. As a by-product of denoising and increased sharpness, artifacts were likely better delineated as these were inherent to the acquired data. However, these artifacts likely did not impede interpretation: 1) ghosting artifacts appeared in the air surrounding the anatomy and 2) pulsation artifacts were largely offset from nerves due to the anterior-to-posterior phase-encoding direction being orthogonal to the predominantly axial course of the nerves. The relative absence of artifact impact on image interpretation is also supported by the fact that nerve assessment was improved with DLRecon.

DLRecon's effect on improvements in interrater agreement was greater for outer epineurium conspicuity than for fascicular architecture. For this study, both readers had dedicated musculoskeletal MRI but variable MR neurography experience, which could possibly explain the improved interrater agreement in outer epineurium conspicuity. We speculate that these differences were attributable to DLRecon's effects on variable image textures: the fascicular architecture being more point-like and the outer epineurium being more edge-like.

Study limitations include a moderate sample size, and varying anatomy and imaging parameters that increased variability in the acquired data. MR neurography, as performed at our institution, frequently involves imaging protocols tailored to each case in order to maximize diagnostic yield (in particular spatial resolution), which invariably results in different sampling matrices even for the same anatomical region. Anecdotally, the effect of DLRecon is even greater at matrices smaller than used in this study; as such, greater improvements could be realized with uniform, albeit low-resolution parameters. Another limitation was that a single contrast (intermediate-weighted) was evaluated, as that was among the most common contrast acquired among the different MRI protocols. MR neurography also uses heavily-T2-weighted fat-suppressed sequences, which were not evaluated in this study. Specifically, the majority of MR neurography scans at our institution employ either 2D multi-echo, Dixon-based FSE or 3D short-tau inversion recovery FSE sequences, both currently incompatible with the DLRecon software. A limitation related to interpretation was that the two readers analyzed the entire image stack rather than scoring single images per nerve. While this likely caused increased scoring variability, this approach was chosen to better reflect standard clinical practice.

One advantageous feature of DLRecon is that it can be applied alongside standard reconstruction and acquisition schemes as it neither alters k-space coverage nor requires a different acquisition type. As edges and details are typically time-sensitive to obtain, due to encoding for image details residing at the outer edges of k-space, we believe MR neurography to be a desirable application for DLRecon. However, we believe DLRecon can be applied to other common musculoskeletal MRI exams, particularly for the detection of chondral and labral abnormalities in the hip and shoulder. In our study, AI improved overall image quality for the same acquired data, but it could instead be used to increase acquired spatial resolution (in-plane and/or through-plane) for the same scan time (e.g., via parallel imaging) or to reduce scan time (fewer acquisitions, higher bandwidth) while maintaining the same resolution. In the near future, we envision DLRecon's application to 3D acquisitions, which may further increase the possibilities of improving image quality as 3D data generally provides higher SNR than 2D. Finally, as DLRecon is

applied to denoise image space dimensions, denoising methods that operate in other dimensions such as echo time [17] and diffusion [18] may be used in combination with DLRecon to further improve image quality.

In conclusion, the results of this study suggest that MR images reconstructed with the novel DLRecon method demonstrate improved outer epineurium and fasicular architecture conspicuity compared to MR images reconstructed with the SOC method. Given that outer epineurium and fasicular architecture consiquity are two key morphological features that are critical to evaluating a nerve injury, these improvements could add diagnostic value to the assessment of clinically suspected peripheral neuropathy. Although DLRecon images had greater pulsation and ghosting artifacts compared to SOC images, they did not affect image interpretation. Given the results of this study, future work will focus on assessing the added clinical value of the DLRecon method in detection of chondral and labral abnormalities in the hip and shoulder and other common musculoskeletal exams.

Declarations

Data Availability: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Acknowledgments: None.

Author Contributions: All authors contributed to the generation and analysis of data in this study. KZ generated the original manuscript draft. ETT, EA, and DBS acquired and reviewed the included figures. BL performed the statistical analysis. All authors reviewed and edited the manuscript.

Competing Interests: The Hospital for Special Surgery, the institution of seven of the authors, has an institutional research agreement with General Electric Healthcare.

References

1. Stewart JD (2003) Peripheral nerve fascicles: Anatomy and clinical relevance. *Muscle Nerve* 28(5):525-541
2. Sunderland S (1951) A classification of peripheral nerve injuries producing loss of function. *Brain* 74(4):491-516
3. Chhabra A, Flammang A, Padua A, Carrino JA, Andreisek G (2014) Magnetic resonance neurography: Technical considerations. *Neuroimaging Clin N Am* 24(1):67-78
4. Soher BJ, Dale BM, Merkle EM (2007) A Review of MR Physics: 3T versus 1.5T. *Magn Reson Imaging Clin N Am* 15(3):277-290
5. Sodickson DK, Griswold MA, Jakob PM, Edelman RR, Manning WJ (1999) Signal-to-noise ratio and signal-to-noise efficiency in SMASH imaging. *Magn Reson Med* 41(5):1009-1022

6. Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P (1999) SENSE: sensitivity encoding for fast MRI. *Magn Reson Med* 42(5):952-962
7. Lustig M, Donoho D, Pauly JM (2007) Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* 58(6):1182-1195
8. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Trans Image Process* 26(7):3142-3155
9. Dong C, Loy CC, He K, Tang X (2014) Learning a Deep Convolutional Network for Image Super-Resolution. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, eds. *Computer Vision – ECCV 2014*. Springer International Publishing 184-199
10. Hammernik K, Klatzer T, Kobler E, et al (2018) Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med* 79(6):3055-3071
11. Lebel RM (2020) Performance characterization of a novel deep learning-based MR image reconstruction pipeline. http://arxiv.org/abs/2008.06559.
12. Kim M, Kim HS, Kim HJ, et al (2021) Thin-slice Pituitary MRI with Deep Learning–based Reconstruction: Diagnostic Performance in a Postoperative Setting. *Radiology* 298(1):114-122
13. van der Velde N, Hassing HC, Bakker BJ, et al (2020) Improvement of late gadolinium enhancement image quality using a deep learning–based reconstruction algorithm and its influence on myocardial scar quantification. *Eur Radiol* doi:10.1007/s00330-020-07461-w (online ahead of print, January 17, 2021)
14. Antun V, Renna F, Poon C, Adcock B, Hansen AC (2020) On instabilities of deep learning in image reconstruction and the potential costs of AI. *Proc Natl Acad Sci* 117(48): 30088-30095
15. Kingma DP, Ba J (2014) Adam: A Method for Stochastic Optimization. http://arxiv.org/abs/1412.6980.
16. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159-174
17. Does MD, Olesen JL, Harkins KD, et al (2019) Evaluation of principal component analysis image denoising on multi-exponential MRI relaxometry. *Magn Reson Med* 81: 3503-3514
18. Sperl JI, Sprenger T, Tan ET, Menzel MI, Hardy CJ, Marinelli L (2017) Model-based denoising in diffusion-weighted imaging using generalized spherical deconvolution. *Magn Reson Med* 78(6):2428-2438

Tables

Table 1. Study Participant Demographics

Characteristic	Value
Total participants	29
Number of males	12
Mean age±standard deviation	49.5±17.8 years
Number of females	17
Mean age±standard deviation	49.4±15.2 years
Total nerves examined	60 nerves
Sciatic nerve	21 (35%)
Tibial nerve	16 (27%)
Median nerve	16 (27%)
Ulnar nerve	7 (12%)
Body parts imaged	
Thigh	13
Knee	4
Lower Leg	3
Elbow	3
Forearm	2
Arm	1
Wrist	1
Hand	1
Ankle	1

Table 2. Distribution of Image Quality Factors Evaluated in SOC-MRIs and DLRecon-MRIs

		Reader 1				Reader 2			
Feature	Grade	DLRecon (N=60)		SOC (N=60)		DLRecon (N=46)		SOC (N=46)	
		n	%	n	%	n	%	n	%
Outer Epineurium Conspicuity (# borders visualized)	0	1	2%	1	2%	0	0%	0	0%
	1	2	3%	4	7%	2	4%	3	7%
	2	3	5%	4	7%	11	24%	17	37%
	3	5	8%	5	8%	17	37%	16	35%
	4	49	82%	46	77%	16	35%	10	22%
Visualization of Fasicular Architecture	poor	7	12%	10	17%	3	7%	6	13%
	average	22	37%	28	47%	16	35%	19	41%
	good	18	30%	14	23%	16	35%	15	33%
	excellent	13	22%	8	13%	11	24%	6	13%
Pulsation Artifact	none	31	52%	30	50%	15	33%	19	41%
	mild	6	10%	16	27%	9	20%	10	22%
	moderate	16	27%	9	15%	11	24%	11	24%
	severe	7	12%	5	8%	11	24%	6	13%
Ghosting Artifact	none	21	35%	26	43%	1	2%	5	11%
	mild	12	20%	14	23%	12	26%	20	43%
	moderate	14	23%	16	27%	22	48%	15	33%
	severe	13	22%	4	7%	11	24%	6	13%
Bulk Motion	none	40	67%	41	68%	43	93%	44	96%
	mild	10	17%	14	23%	1	2%	1	2%
	moderate	8	13%	1	2%	2	4%	1	2%
	severe	2	3%	4	7%	0	0%	0	0%

Table 3. Comparing Odds Ratios (OR) of DLRecon-MRI Scoring Higher than SOC-MRI for Image Quality Features Evaluated by Reader 1 and Reader 2

Feature	Reader 1			Reader 2		
	DLRecon vs SOC			DLRecon vs SOC		
	OR	95% CI	p-value	OR	95% CI	p-value
Outer Epineurium	1.4	(0.6, 3.4)	0.478	1.9	(1.2, 3.1)	0.007
Fasicular Architecture	1.8	(1.0, 3.1)	0.057	1.8	(1.3, 2.5)	<.001
Pulsation Artifact	1.2	(0.9, 1.6)	0.141	1.6	(1.2, 2.2)	0.004
Ghosting Artifact	1.7	(1.3, 2.4)	0.001	2.8	(1.7, 4.6)	<.001
Bulk Motion	1.2	(0.5, 2.6)	0.719	N/A (95% of scores are 0, no variability)		

Table 4. Inter-Reconstruction and Interrater Agreement for Image Quality Features evaluated in SOC- and DLRecon MR Images

Feature	AC (95% cbCI)			
	Inter-reconstruction agreement: SOC vs DLRecon		Interrater Agreement: Reader 1 vs Reader 2	
	Reader 1 (n=60)	**Reader 2 (n=46)**	**SOC (n=46)**	**DLRecon (n=46)**
Outer Epineurium	0.86 (0.75, 0.93)	0.80 (0.63, 0.90)	0.34 (0.06, 0.59)	0.71 (0.42, 0.79)
Fasicular Architecture	0.73 (0.62, 0.82)	0.86 (0.77, 0.94)	0.75 (0.64, 0.85)	0.73 (0.61, 0.86)
Pulsation Artifact	0.87 (0.73, 0.97)	0.83 (0.71, 0.95)	0.86 (0.69, 0.98)	0.76 (0.65, 0.91)
Ghosting Artifact	0.77 (0.63, 0.91)	0.80 (0.70, 0.87)	0.64 (0.44, 0.79)	0.57 (0.35, 0.74)
Bulk Motion	0.77 (0.53, 0.92)	0.98 (0.93, 0.99)	0.84 (0.58, 0.95)	0.86 (0.56, 0.95)

Strength of agreement was interpreted using the following scale: less than 0, poor agreement; 0.00-0.20, slight; 0.21-0.40, fair; 0.41-0.60, moderate; 0.61-0.80, substantial; and 0.81-1.0, almost-perfect agreement.[16] AC= ordinal weighted Gwet's agreement coefficient. cbCI= patient-clustered bootstrap confidence intervals to account for patients who had more than 1 nerve examined.

Figures

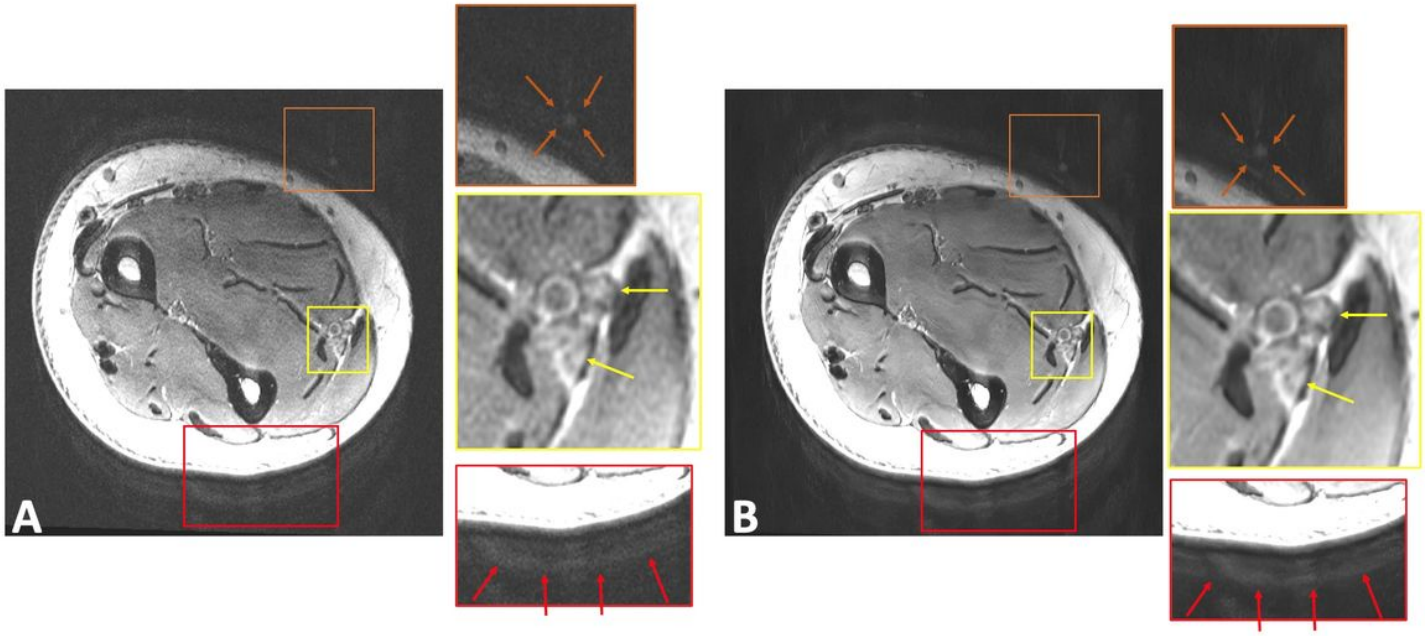

Figure 1

23-year-old female presenting for non-specific median nerve symptoms. Axial FSE intermediate-weighted images processed with DLRecon (B) demonstrated more prominent ghosting artifact (score=2, moderate) and pulsation artifact (score=2, moderate) as compared to the same images processed with the SOC reconstruction method (A, score=1, mild for both ghosting and pulsation artifact), as seen in the magnified insets (arrows, red inset for ghosting; arrows, orange inset for pulsation). However, the increased ghosting and pulsation artifacts did not interfere with nerve conspicuity (yellow inset). The median nerve appeared normal on the exam.

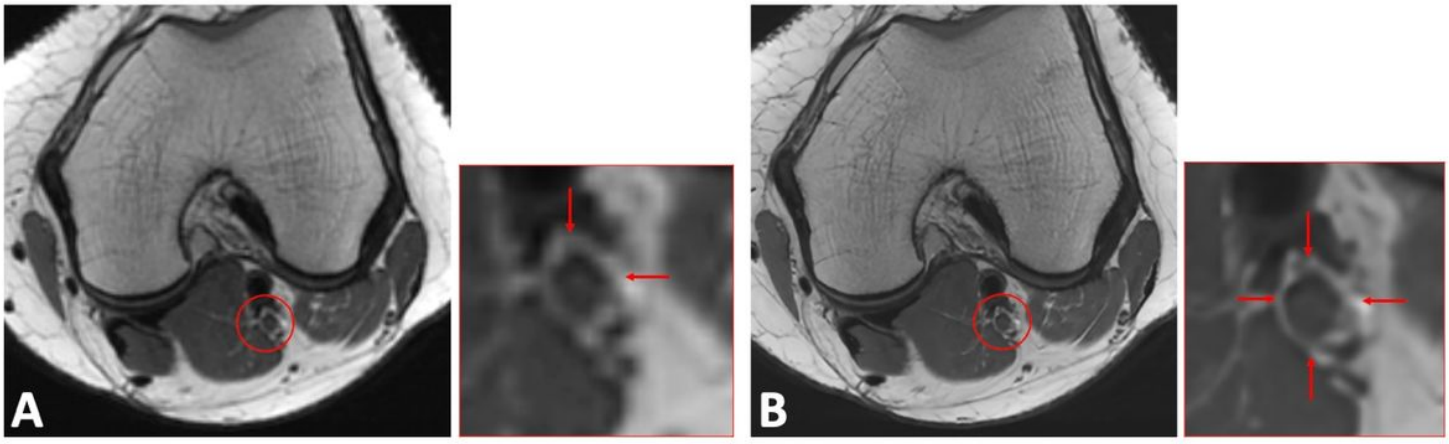

Figure 2

23-year-old female with plantar foot pain. MR neurography was performed of the tibial nerve (circle, inset). The axial intermediate-weighted FSE image at the level of the left knee processed using the DLrecon method (B) exhibits superior outer epineurium conspicuity (4 borders visualized), compared to

the same image processed using the SOC reconstruction method (A, 2 borders visualized). No abnormality of the tibial nerve was identified on this exam.

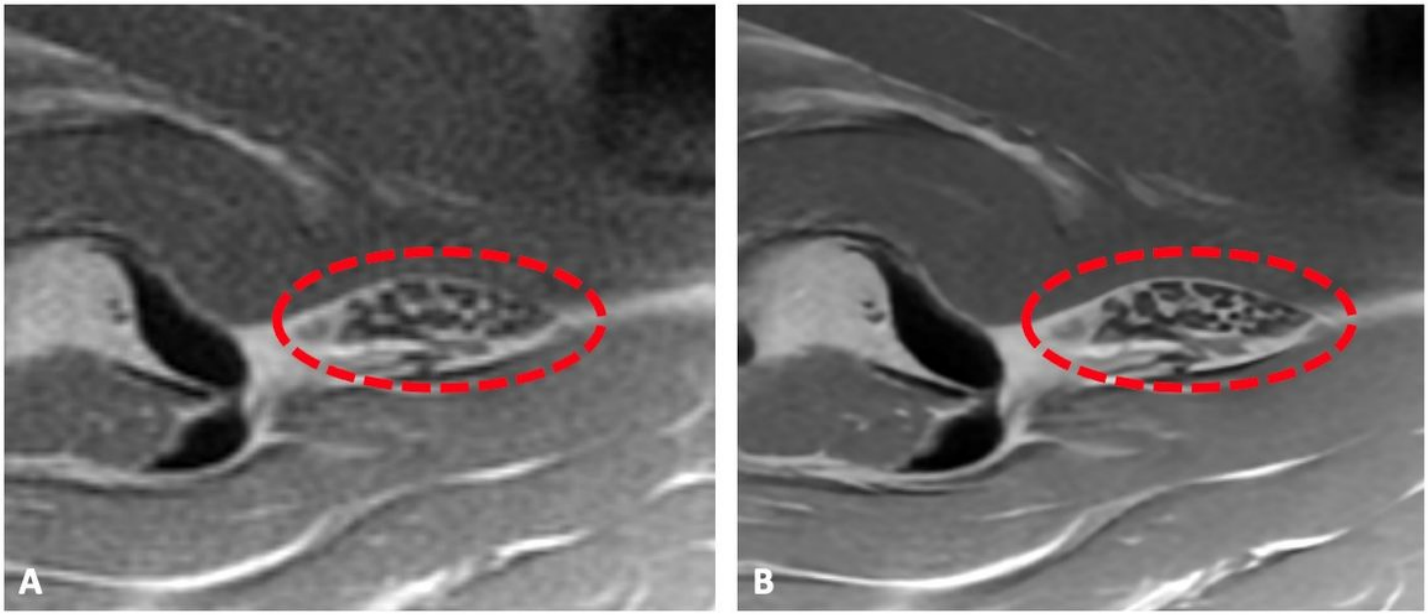

Figure 3

45-year-old male with left sciatic symptoms. Axial intermediate-weighted FSE images within the proximal thigh processed using the DLRecon method (B) exhibited superior visualization of fascicular architecture (score=4, excellent) compared to images processed using the SOC reconstruction method (A, score=2, average). The sciatic nerve appeared normal on this exam.

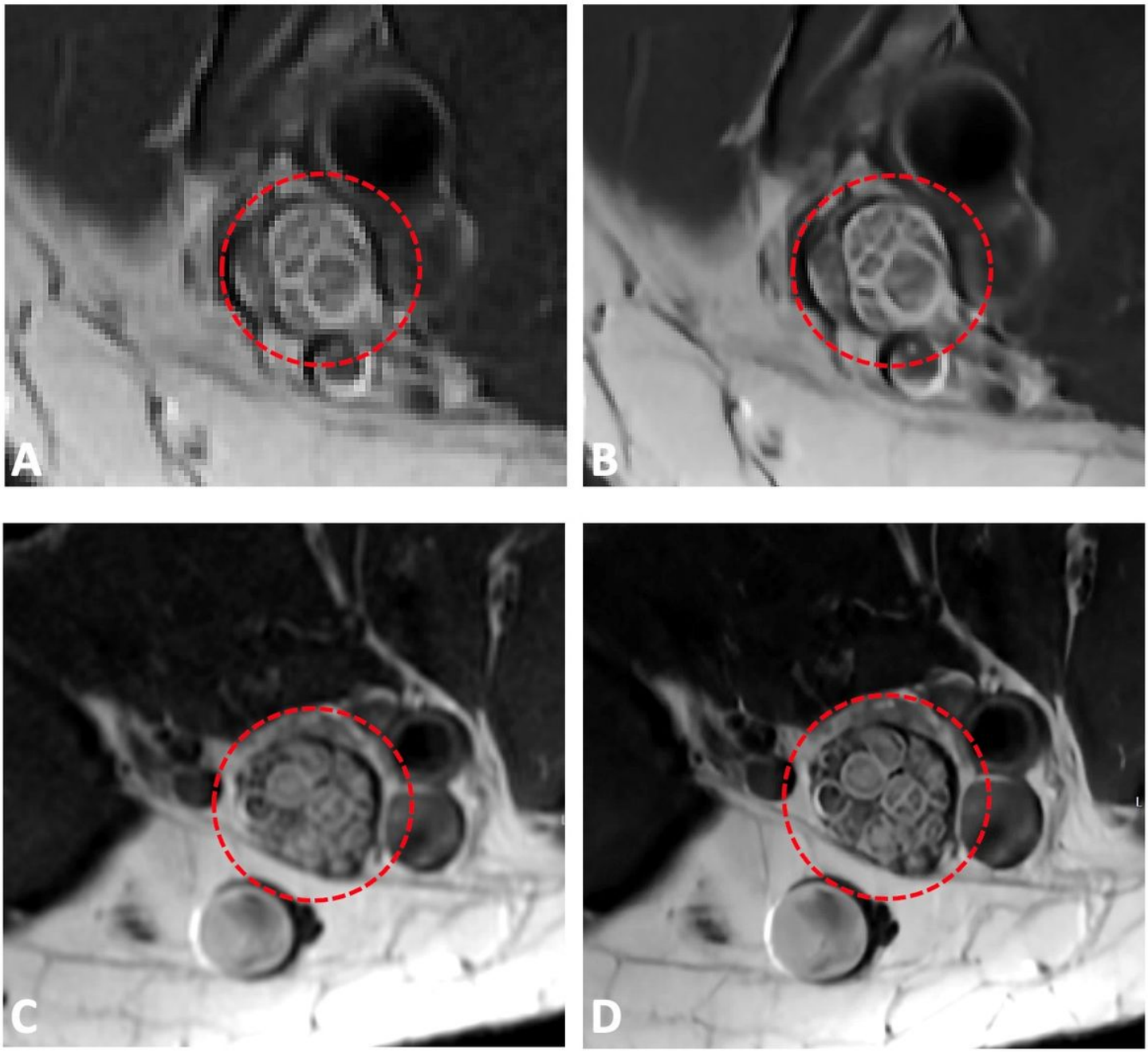

Figure 4

Two patients with spontaneous median motor neuropathy. A 35-year-old female (A and B) and a 24-year-old male (C and D) with spontaneous median motor neuropathy were found to have prominent long-segment fascicular enlargement and signal hyperintensity of their median nerve in the arm by MR neurography. Axial intermediate-weighted FSE images processed with DLRecon (B, D) exhibited improved fascicular architectural detail (score=4, excellent) compared to the same images processed with the SOC reconstruction method (A and C) (score=2, average).