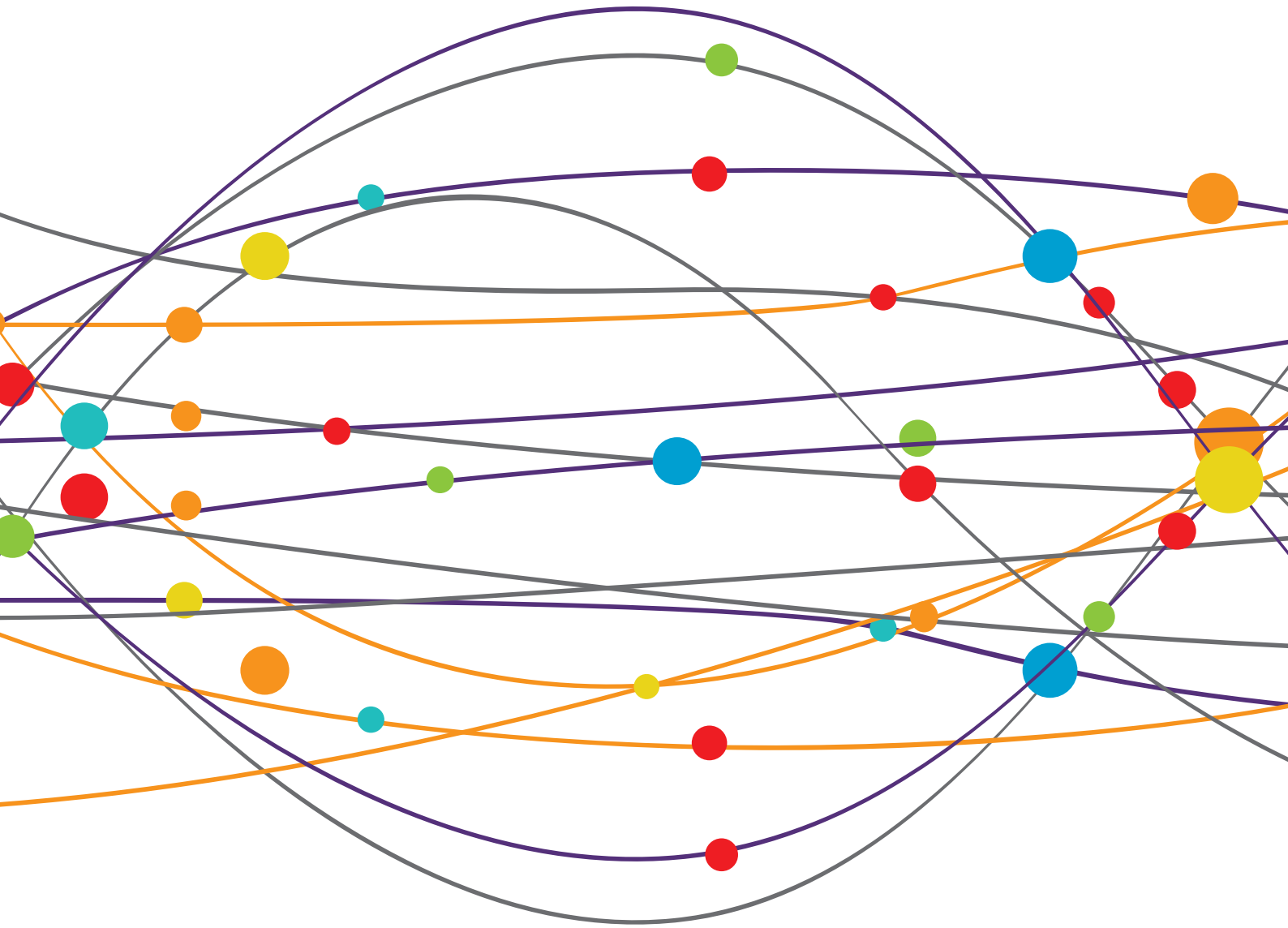
IMAGING OF NEUROMUSCULAR DISEASES

EDITED BY: Massimiliano Filosto, Anna Pichiecchio, Jordi Diaz-Manera and Francesco Santini

PUBLISHED IN: Frontiers in Neurology



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88974-110-6
DOI 10.3389/978-2-88974-110-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

IMAGING OF NEUROMUSCULAR DISEASES

Topic Editors:
Massimiliano Filosto, University of Brescia; NeMO-Brescia Clinical Center for Neuromuscular Diseases, Italy
Anna Pichiecchio, Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy
Jordi Diaz-Manera, University of Newcastle, United Kingdom
Francesco Santini, University of Basel, Switzerland

Citation: Filosto, M., Pichiecchio, A., Diaz-Manera, J., Santini, F., eds. (2022). Imaging of Neuromuscular Diseases. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88974-110-6

Table of Contents

frontiers in Neurology

EDITORIAL
published: 02 December 2021
doi: 10.3389/fneur.2021.814579

Editorial: Imaging of Neuromuscular Diseases

Jordi Diaz-Manera[1], *Anna Pichiecchio*[2,3], *Francesco Santini*[4,5,6] *and Massimiliano Filosto*[7,8]*

[1] John Walton Muscular Dystrophy Research Centre, University of Newcastle, Newcastle upon Tyne, United Kingdom, [2] Istituto di Ricovero e Cura a Carattere Scientifico Mondino Foundation, Pavia, Italy, [3] Brain and Behavioral Sciences Department, University of Pavia, Pavia, Italy, [4] Basel Muscle Magnetic Resonance Imaging (MRI), Department of Biomedical Engineering, University of Basel, Basel, Switzerland, [5] Division of Radiological Physics, Department of Radiology, University Hospital Basel, Basel, Switzerland, [6] Department of Research and Analytic Services, University Hospital Basel, Basel, Switzerland, [7] Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy, [8] NeMO-Neuromuscular OminCenter-Brescia Clinical Center for Neuromuscular Diseases, Brescia, Italy

Keywords: neuromuscular diseases, myopathies, neuropathies, imaging, ultrasound, MRI

Editorial on the Research Topic

Imaging of Neuromuscular Diseases

The heterogeneity of clinical presentation and the rarity of many neuromuscular diseases make it often difficult to achieve the correct diagnosis which is therefore delayed over time.

Recent years have seen a renewed interest in neuromuscular diseases especially due to the emerging innovative therapies, such as molecular therapies, gene therapies, enzyme therapies, already available for some of them such as Duchenne Muscular Dystrophy, Spinal Muscular Atrophy, transthyretin amyloid neuropathy and Pompe disease.

New and exciting therapeutic frontiers await us in the near future and, consequently, early diagnosis has become absolutely necessary for treating patients in the early stages of the disease.

Diagnosis of neuromuscular diseases is historically carried out on the basis of the clinical picture, electromyographic and electroneurographic examination and pathological study by means of a muscle or nerve biopsy.

In recent years, technological development has made available innovative imaging methods that have made it possible to study muscle structures in detail, analyze the severity and distribution of damage and identify specific patterns of damage distribution.

Similarly, ultrasound and MR have become essential tools to support the clinician in diagnosing both peripheral mono- and polyneuropathies (Neurography MR) as well as motor neuron pathology (DTI tractography).

The aim of this Topic Issue was to provide novel original findings on the role of imaging in diagnosing neuromuscular diseases. Several researchers contributed interesting experience and point of views on this subject that is under continuous development.

This topic issue has collected contributions that can be summarized in four macro areas: development of new methods, applications of imaging to muscle diseases, study of peripheral nerve diseases, and evaluation of muscle damage secondary to central nervous system diseases.

NOVEL METHODS

In general, neuromuscular diseases are rare diseases. For this reason, a higher level of standardization is required in order to compare results across different patients, who are often examined at different imaging centers. To achieve this, standardized quantification and evaluation approaches as well as innovative acquisition methods have emerged.

OPEN ACCESS

Edited and reviewed by:
Giovanni Meola,
University of Milan, Italy

***Correspondence:**
Massimiliano Filosto
massimiliano.filosto@unibs.it

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: *13 November 2021*
Accepted: *17 November 2021*
Published: *02 December 2021*

Citation:
Diaz-Manera J, Pichiecchio A, Santini F and Filosto M (2021) Editorial: Imaging of Neuromuscular Diseases. Front. Neurol. 12:814579. doi: 10.3389/fneur.2021.814579

As regards to the acquisition, the functional aspect of the skeletal muscle is acquiring a prominent role in MR imaging. Mazzoli et al. reported a diffusion-sensitive sequence based on oscillating-gradient spin echo (OGSE) instead of the more common pulsed-gradient spin echo (PGSE) for the acquisition of the water diffusion tensor during muscle contraction. The authors demonstrated fewer artifacts and more reliable quantitative values due to the inherent robustness of this gradient pattern to motion.

On the post processing/analysis aspect, Santini et al. proposed an efficient method for the extraction of quantitative T2 values from multi-echo spin-echo images by using an efficient implementation based on GPU processing. By including an external fat fraction measurement into the processing, the authors demonstrated that it is possible to reduce the acquisition protocol to clinically acceptable times. The fitting method was released as an open source software, an important step toward reproducibility.

When specialized acquisitions are not available, Akinci D'Antonoli et al. demonstrated the utility of texture analysis and machine learning methods to assist the diagnosis of cerebral palsy (CP) cases. Texture analysis extracts quantitative features from qualitative images (in this case, in-phase gradient-echo images) and it potentially improves objectivity and reader-dependency. In their paper, the authors demonstrated the superiority of their texture analysis approach with respect to simple fat fraction for the diagnosis of paretic muscles in CP.

All quantitative evaluations of the imaging features in neuromuscular diseases, however, depend on an accurate and reliable segmentation of the muscle groups. This is a challenging task because of deformable geometry, different appearance due to pathological processes, and the limited number of cases available for each site. Ogier et al. gave a detailed overview of the state-of-the-art methods for muscle segmentation including different algorithms and strategies. Atlas-based and conventional image analysis methods are currently being replaced by deep-learning-based methods. While these methods are showing promising results, the availability of large amounts of labeled data is a limit that needs to be addressed for the building of accurate models.

IMAGING APPLIED TO THE STUDY OF MUSCLE DISORDERS

Muscle MR has already demonstrated to be an useful tool for the diagnosis and follow-up of patients with a wide variety of primary muscle disorders including both acquired and inherited myopathies. There are a large number of studies using conventional MR sequences such as T1-weighted and STIR describing the changes in muscle structure of patients at different stages of disease progression. T1w imaging is useful as it identifies fat tissue in the skeletal muscles and allows to describe the so-called patterns of muscle involvement which could guide the diagnosis. But as T1w allows to quantify fat replacement, these data have been successfully used to study correlation with results of muscle function.

Here, Reyes-Leiva et al. described the correlation between the amount of fat replacement in respiratory accessory muscles and the results of the spirometry in a large cohort of patients with late onset Pompe disease, confirming that fat replacement of these muscles could be used as a biomarker of disease progression. In this sense quantification of fat through muscle MR can also be used to classify patients as severely or moderately affected leading to the identification of risk factors that can be associated to one or the other group.

Moore et al. studied a large cohort of patients with dysferlinopathy and identified that performing exercise during adolescence was associated with a more severe fat replacement on the pelvis, thigh and leg muscles and to worse results on muscle function tests. Quantification of muscle fat replacement by means of Dixon sequences has proven to be more powerful to identify changes over time in muscle fat content than semi quantitative methods. Moreover, in most of these studies, muscle function tests are stable over short periods of time although fat content increases as detected by quantitative sequences, suggesting that subtle changes in muscle structure, which are a sign of disease progression, are not always translated into changes in muscle function.

Sheikh et al. included 16 Becker muscular dystrophy patients in 1 year longitudinal study where axial and lower limb muscles were imaged using Dixon MRI and their strength was assessed using Biodex. The authors observed that frat fraction was higher in paraspinal muscles in Becker patients compared to controls and that there was an inverse significant correlation between the fat fraction and muscle strength of the paraspinal muscles. Moreover, they also observed significant differences in fat replacement over 1 year period of time in the paraspinal muscles using Dixon in those patients that were less affected at baseline while there were no differences in their muscle strength suggesting that MRI can be a stronger biomarker for disease progression than muscle function test in clinical trials or natural history studies.

Detection of fibrosis is one of the unmet needs of muscle MRI. Although there have been several attempts to identify fibrotic tissue, so far there is not any reliable MRI sequence able to identify and/or quantify collagen in the skeletal muscles. Murphy et al. successfully used EP3533, which is a novel MRI contrast agent with an affinity to collagen 1 that correlates to *ex vivo* fibrosis quantification, to study response to treatment of mdx mice treated with halofuginone, a well-known anti-fibrotic compound. They were able to observe an increase in fibrotic tissue in the non-treated animals that was statistically higher than the treated group. Histological findings were in agreement with EP35333 imaging. Although further studies in humans are needed, this study opens the door to use contrast agents to detect and quantify fibrotic content in muscles of patients with muscular dystrophies. Other components of the skeletal muscles can also be detected using specific sequences such as glycogen. However, these innovative sequences are generally not available in hospitals.

With the aim to identify indirect changes on muscle MRI conventional sequences of glycogen, Nuñez-Peralta et al. studied a cohort of patients with late onset Pompe disease with

magnetization transfer sequence which analyzed the transfer imaging that exploits the interaction between bulk water protons and protons contained in macromolecules. Authors observed a decreased magnetization transfer ratio (MTR) in skeletal muscles of patients with Pompe disease that correlated with fat content measured using Dixon suggesting that MTR is a good marker of loss of muscle fibers and substitution by fatty tissue. Authors however were not able to see any changes in muscles with low levels of fat replacement, suggesting that if glycogen is accumulated in these muscles this is not inducing any change in MTR.

IMAGING APPLIED TO THE STUDY OF PERIPHERAL NERVOUS SYSTEM

The diagnostic workup of peripheral neuropathies is traditionally based on clinical history, physical examination and electrophysiological studies. In the last few years, novel MR and ultrasound (US) imaging techniques have been developed for studying the peripheral nervous system which have increased the diagnostic rate in this field of neurology.

Ultrasound is very suitable for evaluation of superficial peripheral nerves and MR is especially useful in studying deeply located nerves.

Carpal tunnel syndrome (CTS) is one of the most common upper limb compression neuropathies which is usually diagnosed by clinical history and examination and electrophysiological studies. Ultrasonography is commonly used to confirm diagnosis after electrodiagnostic testing and a median nerve CSA increase at the carpal tunnel inlet is usually observed. Fixed values for the upper limit of normal (ULN) with a broad range of 8.5–15 mm^2 are reported.

In a first study Dubbelink, De Kleermaeker, Beekman et al. showed that the use of a wrist-circumference dependent cut-off value for the CSA of the median nerve at the wrist has a higher sensitivity than using a fixed cut-off of 11 mm^2 or an intraneural flow related cut-off and may be especially useful in patients with a smaller wrist circumference.

In a second study, Dubbelink, De Kleermaeker, Meulstee et al. measured CSA and wrist circumference in a prospective cohort of 253 clinically defined CTS patients and 96 healthy controls and developed an equation for the ULN for CSA by means of univariable regression analysis. They found an augmented diagnostic accuracy of this newly developed equation with a corresponding sensitivity and specificity of 75% compared to a sensitivity of 70% by using a fixed cut-off value of 11 mm^2 and therefore improving diagnostic accuracy of ultrasonography in patients with CTS.

Nerve ultrasound is useful in differential diagnosis of polyneuropathies.

Du et al. evaluated whether ultrasound is suitable for differentiating Transthyretin familial amyloid polyneuropathy (TTR-FAP) and chronic inflammatory demyelinating polyneuropathy (CIDP) because misdiagnosis is frequent as of similar phenotypes. By performing consecutive ultrasonography scanning in six pairs of nerves of bilateral limbs with 30 site and comparing CSAs and CSA variability data in 18 patients with TTR-FAP, 13 patients with CIDP and 14 healthy controls, they showed that both TTR-FAP and CIDP present larger CSAs at most sites of both upper and lower limbs than in control groups but the CSA variability of median nerves is significantly higher in CIDP than in TTR-FAP and control groups with high sensitivity and specificity to differentiate CIDP from TTR-FAP, therefore suggesting that nerve ultrasound can be a potential auxiliary tool to help differentiate the two polyneuropathies.

Imaging can also be a very useful tool in detecting involvement of peripheral nervous system when clinical and electrophysiological assessments are inconclusive.

Kim and Sung studied by MR two cases of neurogenic thoracic outlet syndrome (N-TOS) which is a chronic compressive brachial plexopathy involving the C8 and T1 roots, and/or lower trunk. Although patients showed typical neurological symptoms of N-TOS and structural abnormalities of the thoracic outlet (cervical rib and anomalous first rib), no neurological deficit at the neurologic examination and no abnormalities in the electrophysiological studies were observed. Diagnosis was achieved by CT angiography showing abnormalities of the subclavian artery and root-plexus MR revealing anomalies of nerve root or lower trunk.

Mandelli et al. studied paraspinal muscles by MR in patients with symptomatic Lumbar spinal stenosis in order to compare paraspinal muscle fatty infiltration as assessed using the Goutallier classification vs. quantitative MR measurements. They found that the Goutallier classification is a reliable tool for assessing fatty infiltration of paraspinal muscles in this group of patients and suggest taking body height as a reference for normalization.

CNS DISEASES

Muscles can also be secondly involved in different disorders such as cerebral palsy (CP) or spinal muscular atrophy (SMA).

Muscles from patients with CP are usually spastic and present contractures that limit the range of motion. Weidensteiner et al. assessed by MR the effect of botulinum toxin A (BTX) over time on calf muscle properties in pediatric CP patients. BTX induced increase in extracellular space and a simultaneous decrease of muscle fiber diameter and MRI showed limited spatial distribution of these BTX-induced effects, being a promising non-invasive tool for future studies in order to test BTX treatment protocols.

On the other hand, in muscles of patients affected by SMA3 treated with nusinersen Savini et al. found a progression of fat fraction (FF) in thigh muscles and a concurrent slight reduction of water T2 over time, despite therapy. Current advances in morphometric MRI development also allow gray (GM) and white matter (WM) quantification in the spinal cord (SC), which may help in improving the *in vivo* characterization of neurodegenerative SC diseases or lower motor neuron disorders such as SMA. These advanced MR techniques have been applied together with muscle quantitative MR in the pilot study of Savini et al. on three adult SMA3 patients under

treatment with nusinersen to characterize SC volumes and microstructure (GM and ventral horns) as well. The authors demonstrated that quantitative cervical SC MR sequences derived from the consensus acquisition protocol produced by a consortium of SC researchers (1) including 3D T1- and T2 weighted (w), DWI, MT and T2*-w sequences were able to capture microstructural changes induced by SMA *in vivo* and are a candidate methodology for monitoring the effects of this treatment.

Different SC imaging approaches are also appearing on the horizon, such as radially sampled averaged magnetization inversion recovery acquisition (rAMIRA), which is a novel approach to perform SC imaging in clinical settings with favorable contrast. These approaches need to be tested on healthy subjects (HS) to understand the sources of inter-subject variability in different SC levels as well as their relation to age and sex to facilitate the detection of pathology-associated changes. Kesenheimer et al. identified effective normalization strategies for inter-subject variability reduction in total SC area and SC GM area on a broad aged cohort of HS using rAMIRA.

Taken together, the papers collected in this Issue present several interesting and original findings which may contribute to the knowledge on the role of imaging in neuromuscular disorders and offer novel perspectives for future studies and clinical applications.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

We are grateful to all authors that contributed to the success of the Topic Issue, allowing us to provide the readers an updated overview on this matter. We wish also to thank all the reviewers that contributed with their knowledge and useful advice. Finally, thanks to the Editorial Office that with expertise and in a professional manner helped us in the management of this collection.

REFERENCES

1. Cohen-Adad J, Alonso-Ortiz E, Abramovic M, Arneitz C, Atcheson N, Barlow L, et al. Generic acquisition protocol for quantitative MRI of the spinal cord. *Nat Protocol.* (2021) 16:4611–32. doi: 10.1038/s41596-021-00588-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 25 September 2020
doi: 10.3389/fneur.2020.577052

Augmented Diagnostic Accuracy of Ultrasonography for Diagnosing Carpal Tunnel Syndrome Using an Optimised Wrist Circumference-Dependent Cross-Sectional Area Equation

Tom B. G. Olde Dubbelink[1], Floriaan G. C. M. De Kleermaeker[2], Jan Meulstee[1], Ronald H. M. A. Bartels[3], Franka Claes[4] and Wim I. M. Verhagen[1]*

[1] Department of Neurology, Canisius-Wilhelmina Hospital, Nijmegen, Netherlands, [2] Department of Neurology, VieCuri Hospital, Venlo, Netherlands, [3] Department of Neurosurgery, Radboud University Medical Centre, Nijmegen, Netherlands, [4] Department of Neurology, Vlietland Hospital, Schiedam, Netherlands

OPEN ACCESS

Edited by:
Francesco Santini,
University of Basel, Switzerland

Reviewed by:
Anna Hirschmann,
University of Basel, Switzerland
Eugenio Belatti,
University of Perugia, Italy

***Correspondence:**
Tom B. G. Olde Dubbelink
t.oldedubbelink@cwz.nl

Specialty section:
This article was submitted to Applied Neuroimaging, a section of the journal Frontiers in Neurology

Received: 28 June 2020
Accepted: 20 August 2020
Published: 25 September 2020

Citation:
Olde Dubbelink TBG, De Kleermaeker FGCM, Meulstee J, Bartels RHMA, Claes F and Verhagen WIM (2020) Augmented Diagnostic Accuracy of Ultrasonography for Diagnosing Carpal Tunnel Syndrome Using an Optimised Wrist Circumference-Dependent Cross-Sectional Area Equation. Front. Neurol. 11:577052. doi: 10.3389/fneur.2020.577052

Introduction: In diagnosing carpal tunnel syndrome (CTS) there is no consensus about the upper limit of normal (ULN) of the cross-sectional area (CSA) of the median nerve at the carpal tunnel inlet. A previous study showed wrist circumference is the most important independent predictor for the ULN. In this study we optimised a wrist circumference-dependent ULN equation for optimal diagnostic accuracy and compared it to the generally used fixed ULN of 11 mm^2.

Methods: CSA and wrist circumference were measured in a prospective cohort of 253 patients (clinically defined CTS) and 96 healthy controls. An equation for the ULN for CSA was developed by means of univariable regression analysis. We calculated *z*-scores for all patients and healthy controls, and analysed these scores in a ROC curve and a decision plot. Sensitivity and specificity were determined and compared to fixed ULN values.

Results: We found augmented diagnostic accuracy of our newly developed equation $y = 0.88 * x - 4.0$, where y = the ULN of the CSA and x = wrist circumference. This equation has a corresponding sensitivity and specificity of 75% compared to a sensitivity of 70% while using a fixed cut-off value of 11 mm^2 ($p = 0.015$).

Conclusion: Optimising the regression equation for wrist circumference-dependent ULN cross-sectional area of the median nerve at the wrist inlet might improve diagnostic accuracy of ultrasonography in patients with carpal tunnel syndrome and seems to be more accurate than using fixed cut-off values.

Keywords: carpal tunnel syndrome (CTS), ultrasonography, diagnostics, cross-sectional area (CSA), median nerve, wrist circumference

INTRODUCTION

Carpal tunnel syndrome (CTS) can be diagnosed by taking accurate medical history in combination with clinical assessment (1). Ultrasonography (US) is the most commonly used test after electrodiagnostic testing (EDX), in confirming the clinical diagnosis of CTS. Especially when surgical decompression is considered, EDX or ultrasonography assessment is performed for confirming the diagnosis. The AAOS Clinical Guidelines recommend EDX testing for CTS patients when surgery is being considered (2) while the Dutch consensus CTS guideline states no additional studies are needed in case of classical CTS (3). An earlier study showed that only a minority of surgeons would perform surgery without electrodiagnostic confirmation of CTS (4).

For ultrasonography, alteration of the shape of the median nerve is evaluated and an enlarged cross-sectional area (CSA) of the median nerve at the carpel tunnel is frequently used to confirm CTS diagnosis (5). Currently, using ultrasonographic evaluation of increase in size of the median nerve at the carpal tunnel, fixed values for the upper limit of normal (ULN) with a broad range of 8.5–15 mm^2 are reported (6–14). This broad range of the normal values may be affected by morphometric factors, as well as age and sex as described in the literature (15–18). An earlier study reported similar diagnostic accuracy of sonography to that for EDX studies (11). Because of comparable sensitivity and patient-friendliness, ultrasonography is recommended as the first line diagnostic test for CTS in The Netherlands.

There is, however, no consensus about the upper limit of normal of the CSA of the median nerve at the carpal tunnel inlet. We previously showed a strong correlation between wrist circumference and CSA of the median nerve at the carpal tunnel inlet in subjects without signs or symptoms of carpal tunnel syndrome (15). Furthermore, we developed an equation for the ULN of the CSA which has a relatively low sensitivity (53.4%) but a very high specificity (95%) (19). It is the low sensitivity that hampers the clinical applicability of this equation. We hypothesize that, by optimising the sensitivity and specificity of this equation, an individualised upper limit of normal of the CSA based on wrist-circumference has an higher sensitivity than a fixed upper limit of normal does. We analysed a decision plot based on a receiver operating characteristic (ROC) curve of the healthy controls to augment diagnostic accuracy and we compared this to the generally used fixed ULN of 11 mm^2 (11).

MATERIALS AND METHODS

Methods

We prospectively enrolled patients and healthy control subjects in this observational study. We obtained written informed consent from each patient and healthy control. Approval from the local Medical Ethics Committee was obtained.

Abbreviations: CTS, carpal tunnel syndrome; CSA, cross-sectional area; ULN, upper limit of normal; US, ultrasonography.

Study Population and Sonography Assessment

We recruited 96 healthy control subjects without signs and symptoms of CTS. All 96 healthy controls underwent medical history taking and physical examination (WV, FC). Controls with a history of diabetes mellitus, rheumatoid arthritis, wrist trauma or BMI >35 kg/m^2 were excluded. Controls with bifid median nerves were excludes as well. Both wrists were measured, we randomly included only one wrist and CSA because we did not find any differences in earlier studies (15).

A total of 253 clinically defined CTS patients were included if they had pain and/or paraesthesia in the territory innervated by the median nerve. Two or more of the following clinical CTS criteria had to be fulfilled: (1) nocturnal paraesthesia, (2) aggravation of paraesthesia by activities such as driving a car, riding a bike, holding a book, or holding a telephone, (3) positive Flick sign (paraesthesia relieved by shaking the affected hand). For patients with bilateral CTS only the most severely affected hand was included.

Exclusion criteria were age under 18; history or clinical signs of polyneuropathy or known hereditary neuropathy with liability to pressure palsies; previous trauma or surgery to the wrist; history of rheumatoid arthritis; diabetes mellitus; thyroid disease; alcoholism; arthrosis of the wrist; pregnancy; severe atrophy of the abductor pollicis brevis muscle; bifid median nerve or significant language barrier.

We measured weight, height and wrist circumference at the level of the distal wrist crease using plastic measuring tape. CSA of all subjects was measured at the inlet of the carpal tunnel (Philips Diagnostic Ultrasound System model iU22, 5–17-MHz linear transducer) using the direct trace method. Electrodiagnostic technicians took the measurements, and patients underwent US and EDX according to the protocol of our previous study (15, 20). The used US parameters were: frequency 17 MHz, acoustic power 100%, dynamic range 77 dB, deepness 1.5 cm, focus position 2 cm, gain 80. US was performed by two experienced US technicians, EDX studies by a clinical neurophysiologist (JM). The cross-sectional area of the median nerve was determined by outlining the nerve contour using the inner margin of the hyperechoic rim. The CSA was calculated by the area measurement software (continuous contour trace) of the ultrasound system, rounding all measurements to the nearest 0.01 cm^2. The mean of three separate measurements was taken as CSA at the inlet of the carpal tunnel.

Z-score

We calculated a Z-score specific for an individual wrist circumference according to:

$$z - \text{score} = (X - \mu)/\sigma = (\text{CSA measured} - \text{CSA expected from wrist circumference})/\text{standard deviation}.$$

X being the mean of three actual CSA measurements, μ the expected CSA calculated from the wrist circumference and σ the standard deviation.

Ergo:

$$\text{z-score} * \text{standard deviation} = (\text{CSA measured} - \text{CSA expected from wrist circumference}).$$

This enables us to generate a "new" upper limit of normal based on wrist circumference (maximum CSA expected from wrist circumference).

Statistics

Statistical analysis was performed using SPSS Statistics 26.0. Baseline characteristic for healthy controls and patients were described as mean ± SD and frequency (%). Unpaired *T*-tests were used for continuous variables with normal distribution, and the Mann–Whitney test in case of non-normal distribution for group comparisons of baseline data. McNemar's test was used for paired categorical data. We used univariable regression analyses to create equations for the ULN for CSA. We used *z*-scores, a receiver operating characteristic (ROC) curve and a decision plot to develop a new equation with optimal diagnostic accuracy. Normal distribution of data was assessed visually by plotting a histogram, using a Q–Q plot and the Kolmogorov–Smirnov test. The level of significance was set at 0.05 for all analyses.

RESULTS

In **Table 1** the characteristics of the included healthy controls and patients are presented. The data of the 96 healthy controls were normally distributed. In the healthy controls the mean CSA was significantly smaller in women ($n = 49$); 8.8 mm^2 (SD 1.9), compared to 10.2 mm^2 (SD 1.9) in men ($n = 47$; $p = 0.001$).

253 patients were consecutively enrolled in our study. In the patient group mean CSA was 13.4 mm^2 (SD 4.4) in women ($n = 202$) and 14.0 mm^2 (SD 4.1) in men ($n = 51$) with no statistically significant difference ($p = 0.132$).

Figure 1 is a scatter plot showing the results of the regression analysis of the 96 healthy controls for determining the upper limit of normal of the CSA depending on wrist circumference.

TABLE 1 | Healthy controls and patient characteristics (standard deviation between brackets).

	Healthy controls	Patients	*p*
Participants (*n*)	96	253	
Men/women	47/49	51/202	<0.001
Mean age (y)	44.6 (±11.4)	47.1 (±10.9)	0.060
Median age (y)	46	48	
Left/right	48/48	149/104	
Mean height (cm)	175.8 (±9.0)	167.3 (±7.9)	<0.001
Mean weight (kg)	77.3 (±13.5)	76.6 (±15.6)	0.692
Mean BMI	25 (±3.6)	27.3 (±4.9)	<0.001
Mean wrist circumference (cm)	16.8 (±1.4)	16.6 (±1.2)	0.293
Mean CSA (mm^2)	9.5 (±2.0)	13.5 (±4.4)	<0.001

The values for expected CSA, μ, were calculated by filling in the wrist circumference in the regression equation for median values. This was used to determine the *Z*-scores for all patients and healthy controls. The difference between the upper limit of normal (upper grey dashed) line and the median value (red) line is 3.25 and equals 1.96 times the standard deviation.

By plotting the earlier mentioned *Z*-scores of all individuals (patients and healthy controls) in a receiver operating characteristic (ROC) curve we can determine the optimal coefficients for the formula for the ULN of wrist circumference-dependent CSA (**Figure 2**). To help determine the optimum decision level (maximum number of CTS patients correctly diagnosed as positive by ultrasonography in relation to maximum number of healthy controls correctly diagnosed as negative) we made a decision plot (**Figure 3**).

The discrimination between healthy controls and patients with clinically defined CTS is optimal when using a *z*-score of 0.75, as can be seen in **Figure 3**. Rewriting and filling in:

$$\text{z-score} * \text{standard deviation} = (\text{CSA measured} - \text{maximum CSA expected from wrist circumference}),$$

results in

$$\begin{aligned}\text{maximum CSA expected from wrist circumference} &= \text{CSA measured} - (0.75 * 1.66) \\ &= \text{CSA measured} - 1.24.\end{aligned}$$

Accordingly, the maximum wrist circumference related CSA is 1.24 higher than previously calculated while using $y = 0.88 * x - 5.25$. This means that we have to add 1.24 for each y. As a result, and taking into account rounding errors, the optimised equation becomes $y = 0.88 * x - 4.0$. At this *z*-value, the corresponding sensitivity is 75% and the specificity for this optimised equation is 75%.

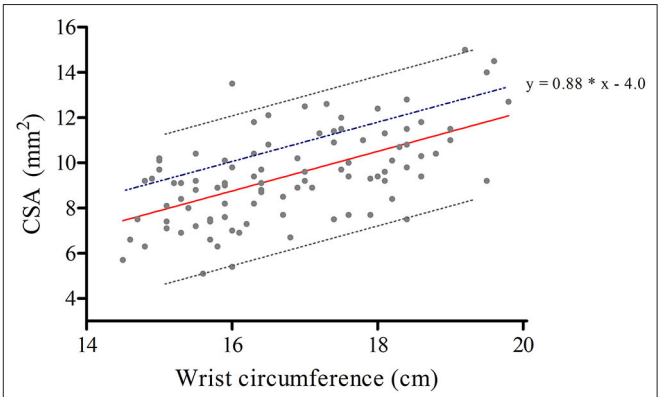


FIGURE 1 | Regression analysis of the 96 healthy controls for determining the upper limit of normal (ULN) of the cross-sectional area depending on wrist circumference. Regression equations for median values (0.88 * x −5.25, red line), and for the upper limit of normal (0.88 * x −2.0, grey upper dashed line); x being the wrist circumference. The blue dashed line represents our optimised equation for the new ULN (y = 0.88 * x −4.0), adding 1.24 for each y.

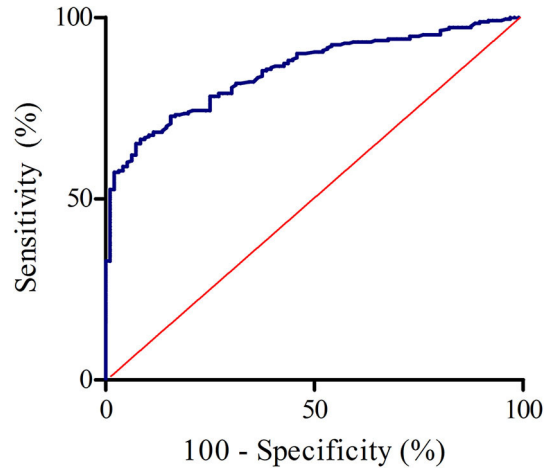


FIGURE 2 | ROC curve of the calculated *z*-scores (healthy controls and clinically defined carpal tunnel syndrome patients). AUC = 0.854, 95% confidence interval: 0.815–0.893.

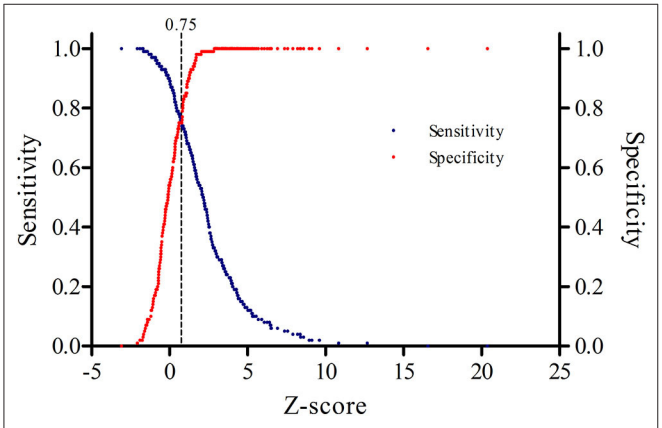


FIGURE 3 | Decision plot showing that a *z*-score of 0.75 (dotted vertical line) represents the optimal cut-off point for the highest sensitivity and specificity.

When applying the wrist circumference-dependent CSA equation and the aforementioned fixed cut-off value in the group of patients, we found the results as presented in **Figure 4**. Abnormal US results were found in 177 (70.0%) patients while using a fixed cut-off value of 11 mm^2 compared to 190 (75.1%) patients when applying our optimised equation. The ultrasound was considered abnormal most often in the wrist circumference-dependent upper limit of normal formula group, with a statistically significant difference compared to the general fixed ULN ($p = 0.015$).

Furthermore, 54/253 (21.3%) of the clinically defined CTS patients had normal EDX results. Eighteen of these 54 (33.3%) patients had an ULN >11 mm^2. Using our new equation (y = 0.88 * x −4.0), 23 of the 54 (42.6%) CTS patients with normal EDX results, would have an abnormal US result.

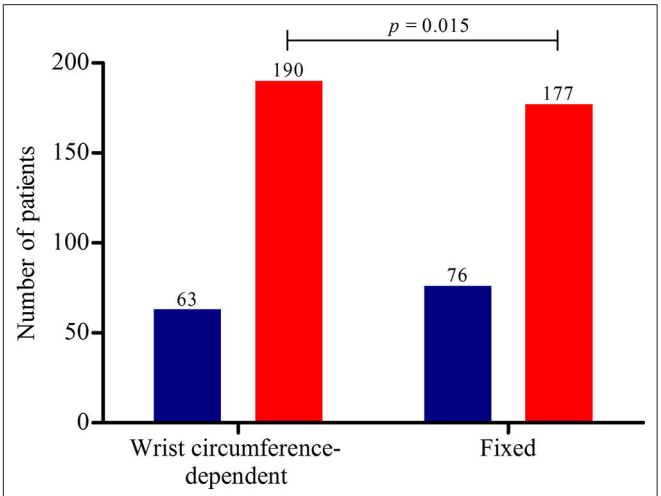


FIGURE 4 | Ultrasound results of the 253 patients. The red bars represent the number of patients with abnormal ultrasonography results and the blue bars normal ultrasonography results. Using a wrist-circumference dependent upper limit of normal (ULN), 190 patients (75.1%) had abnormal US results vs. 177 patients (70.0%) using a fixed ULN of the cross-sectional area of the median nerve of 11 mm^2 ($p = 0.015$). The mean CSA of these 190 and 177 patients with abnormal US results were 14.9 (±4.1) and 15.2 (±4.1), respectively, without a statistically significant difference ($p = 0.201$).

DISCUSSION

We found that our new equation, y = 0.88 * x −4.0 (x = wrist circumference in centimetres), has a corresponding sensitivity of 75% and a specificity of 75%. As presented in **Figure 4**, the sensitivity of our optimised equation is higher than this fixed cut-off value (75.1 vs. 70.0%, $p = 0.015$). By using the data from the ROC approach, increasing sensitivity and decreasing the specificity, it seems that our simple optimised equation is more valuable in daily practice when determining the upper limit of normal of the CSA at the wrist inlet and confirming the clinical suspicion of CTS. Accordingly, wrist circumference-dependent CSA upper limits of normal could be used for better diagnostic accuracy.

Diagnostic accuracy of US in CTS patients depends on the upper limit of normal of the measured CSA of the median nerve at the carpal tunnel inlet. We previously showed, by performing multivariable linear regression, that wrist circumference is the most important independent predicting factor for CSA (15), accounting for 37% of the variation in this parameter (19). We also showed an equation for the ULN of the CSA with a relatively low sensitivity but a very high specificity. In this study we improved the diagnostic accuracy of this equation in order to use US in CTS patients as a screening test. The sensitivity and specificity are rather low at 75% but in line with literature where sensitivity ranges mostly from 70–88% and specificity from 57–97% (21, 22). Earlier studies showed high positive predictive values for ultrasonography: if ultrasonography is abnormal, EDX studies were abnormal in 96.7–100% (20, 23). In daily practice, taking in account the lower

costs of ultrasonography and patient-friendliness, we would suggest to perform ultrasonography as the first diagnostic test in conforming the diagnosis CTS. If ultrasonography is normal, EDX studies could be used as second diagnostic test if clinical suspicion of CTS still exists.

The additional value of ultrasonography compared to EDX studies includes detection of anomalies such as structural abnormalities at the wrist, bifid median nerves and persistent median arteries (24). As stated before US is less time consuming and more comfortable (pain-free) (25). However, not all patients with CTS do have an enlarged median nerve, maybe in part due to a short duration of symptoms (11). EDX studies can quantify nerve damage, have high sensitivity (26) and can be used to differentiate CTS from more proximal median nerve neuropathy or other conditions, for example a C6/C7 radiculopathy.

There are several limitations to our study that should be addressed. First of all, we included CTS patients with sensory symptoms in the fifth finger. We ruled out an ulnar neuropathy in these patients by medical history only, not by nerve conduction studies. However, as stated in a previous study, symptoms in the fifth finger are often reported by CTS patients, and treatment outcome does not differ compared to CTS patients with a classic presentation (27). Secondly, the electrodiagnostic technicians who performed ultrasonography investigating the healthy controls were not blinded. They may have anticipated to find no enlargement of the median nerve and this possibly influenced the measurements. The controls did not have any signs or symptoms of CTS; however, this does not always imply a normal ultrasonography result. This may have influenced the measurements the other way around. We did not calculate intra- and interobserver agreement of the CSA measurements of the median nerves in this study but we found good agreement in an earlier study (15). Furthermore, we excluded healthy controls with a BMI $> 35 kg/m^2$ for a representative univariable regression analysis. We hypothesised that including obese controls could result in a higher ULN of the CSA, leading to more false negatives. 14 of the 253 (5.5%) patients had a BMI >35 kg/m^2 with the expected significantly higher wrist circumference (mean 17.5, $p = 0.019$) but non-significantly higher CSA (mean 14.9, $p = 0.397$) compared to the other patients. In addition, we excluded patients and healthy controls with bifid median nerves so our equation is not valid for these patients.

In the future, evaluating our optimised equation in populations with other morphometric features and less specific populations (not fulfilling all criteria that we used, but a clinical suspicion of CTS) would be interesting, as well as comparing this equation with other parameters, for example an intraneural flow related upper limit of normal (28). In literature, numerous possible ultrasonography parameters have been investigated. The cross-sectional area of the median nerve remains the best single criterion (5) and seems to be related with neurophysiological severity (29). The diagnostic value of wrist median nerve CSA vs. wrist-to-forearm ratio showed inconsistent results in literature (30, 31). A cross-sectional study published in 2019 by Chang et al. suggests the ulnar nerve compared to the median nerve at wrist level could serve as internal control by using median-to-ulnar-nerve difference instead of the median-to-ulnar-nerve ratio (32).

To conclude, this study shows that optimising the regression equation for wrist circumference-dependent ULN cross-sectional area of the median nerve at the wrist by fine tuning its coefficients by ROC curve analysis inlet might improve diagnostic accuracy of ultrasonography in patients with carpal tunnel syndrome and seems to be more accurate than using fixed cut-off values.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Offline dataset. SPSS dataset are available upon request. Requests to access these datasets should be directed to Tom B. G. Olde Dubbelink, t.oldedubbelink@cwz.nl.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Commissie Mensgebonden Onderzoek regio Arnhem-Nijmegen. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TO, FK, JM, FC, and WV contributed conception and design of the study. TO and FK organised the database and performed statistical analysis. TO wrote the first draft of the manuscript. TO, FK, JM, WV, and RB wrote sections of the manuscript. All authors contributed to the manuscript revision and read and approved the final version.

ACKNOWLEDGMENTS

We thank all participants and our electrodiagnostic technicians for their contributions to this study.

REFERENCES

1. Padua L, Coraci D, Erra C, Pazzaglia C, Paolasso I, Loreti C, et al. Carpal tunnel syndrome: clinical features, diagnosis, and management. *Lancet Neurol.* (2016) 15:1273–84. doi: 10.1016/S1474-4422(16)30231-9
2. Keith MW, Masear V, Chung K, Maupin K, Andary M, Amadio P, et al. American academy of orthopaedic surgeons clinical practice guideline on: diagnosis of carpal tunnel syndrome. *J Bone Joint Surg Am.* (2009) 91:10:2478–9. doi: 10.2106/JBJS.I.00643
3. Nederlandse Vereniging voor Neurologie. *Richtlijn Carpaletunnelsyndroom.* (2017) Available online at: https://richtlijnendatabase.nl/richtlijn/carpaletunnelsyndroom_cts/startpagina_-_carpaletunnelsyndroom_cts.html (accessed April 27, 2020).
4. Claes F, Verhagen WIM, Meulstee J. Current practice in the use of nerve conduction studies in carpal tunnel syndrome by surgeons in the Netherlands. *J Hand Surg Eur.* (2007) 32:663–7. doi: 10.1016/J.JHSE.2007.09.007
5. Tai TW, Wu CY, Su FC, Chern TC, Jou IM. Ultrasonography for diagnosing carpal tunnel syndrome: a meta-analysis of

diagnostic test accuracy. *Ultrasound Med Biol.* (2012) 38:7:1121–8. doi: 10.1016/j.ultrasmedbio.2012.02.026
6. Mohammadi A, Afshar A, Etemadi A, Masoudi S, Baghizadeh A. Diagnostic value of cross-261 sectional area of median nerve in grading severity of carpal tunnel syndrome. *Arch Iran Med.* (2010) 13:516–21.
7. Duncan I, Sullivan P, Lomas F. Sonography in the diagnosis of carpal tunnel syndrome. *Am J Roentgenol.* (1999) 173:681–4. doi: 10.2214/ajr.173.3.10470903
8. Fowler JR, Munsch M, Tosti R, Hagberg WC, Imbriglia JE. Comparison of ultrasound and electrodiagnostic testing for diagnosis of carpal tunnel syndrome: study using a validated clinical tool as the reference standard. *J Bone Joint Surg Am.* (2014) 96:e148. doi: 10.2106/JBJS.M.01250
9. Swen WAA, Jacobs JWG, Bussemaker FEAM, De Waard JWD, Bijlsma JWJ. Carpal tunnel sonography by the rheumatologist versus nerve conduction study by the neurologist. *J Rheumatol.* (2001) 28:1:62–69.
10. Deniz FE, Öksüz E, Sarikaya B, Kurt S, Erkorkmaz Ü, Ulusoy H, et al. Comparison of the diagnostic utility of electromyography, ultrasonography, computed tomography, and magnetic resonance imaging in idiopathic carpal tunnel syndrome determined by clinical findings. *Neurosurgery.* (2012) 70:610–16. doi: 10.1227/NEU.0b013e318233868f
11. Visser LH, Smidt MH, Lee ML. High-resolution sonography versus EMG in the diagnosis of carpal tunnel syndrome. *J Neurol Neurosurg Psychiatry.* (2008) 79:63–67. doi: 10.1136/jnnp.2007.115337
12. Nakamichi KI, Tachibana S. Ultrasonographic measurement of median nerve cross-sectional area in idiopathic carpal tunnel syndrome: diagnostic accuracy. *Muscle Nerve.* (2002) 26:798–803. doi: 10.1002/mus.10276
13. Hobson-Webb LD, Padua L. Median nerve ultrasonography in carpal tunnel syndrome: findings from two laboratories. *Muscle Nerve.* (2009) 40:94–7. doi: 10.1002/mus.21286
14. Lee D, Van Holsbeeck MT, Janevski PK, Ganos DL, Ditmars DM, Darian VB. Diagnosis of carpal tunnel syndrome: ultrasound versus electromyography. *Radiol Clin North Am.* (1999) 37:859–72. doi: 10.1016/S0033-8389(05)70132-9
15. Claes F, Meulstee J, Claessen-Oude Luttikhuis TTM, Huygen PLM, Verhagen WIM. Usefulness of additional measurements of the median nerve with ultrasonography. *Neurol Sci.* (2010) 31:721–725. doi: 10.1007/s10072-010-0258-9
16. Zaidman CM, Al-Lozi M, Pestronk A. Peripheral nerve size in normals and patients with polyneuropathy: an ultrasound study. *Muscle and Nerve.* (2009) 40:960–6. doi: 10.1002/mus.21431
17. Won SJ, Kim BJ, Park KS, Yoon JS, Choi H. Reference values for nerve ultrasonography in the upper extremity. *Muscle Nerve.* (2013) 47:864–71. doi: 10.1002/mus.23691
18. Cartwright MS, Shin HW, Passmore LV, Walker FO. Ultrasonographic reference values for assessing the normal median nerve in adults. *J Neuroimaging.* (2009) 19:47–51. doi: 10.1111/j.1552-6569.2008.00256.x
19. De Kleermaeker FGCM, Meulstee J, Verhagen WIM. The controversy of the normal values of ultrasonography in carpal tunnel syndrome: diagnostic accuracy of wrist-dependent CSA revisited. *Neurol Sci.* (2019) 40:1041–7. doi: 10.1007/s10072-019-03756-z
20. Claes F, Kasius KM, Meulstee J, Verhagen WIM. Comparing a new ultrasound approach with electrodiagnostic studies to confirm clinically defined carpal tunnel syndrome: a prospective, blinded study. *Am J Phys Med Rehabil.* (2013) 92:1005–11. doi: 10.1097/PHM.0b013e31829b4bd8
21. Beekman R, Visser LH. Sonography in the diagnosis of carpal tunnel syndrome: a critical review of the literature. *Muscle Nerve.* (2003) 27:26–33. doi: 10.1002/mus.10227
22. Roll SC, Case-Smith J, Evans KD. Diagnostic accuracy of ultrasonography vs. Electromyography in carpal tunnel syndrome: a systematic review of literature. *Ultrasound Med Biol.* (2011) 37:1539–53. doi: 10.1016/j.ultrasmedbio.2011.06.011
23. Pastare D, Therimadasamy AK, Lee E, Wilder-Smith EP. Sonography versus nerve conduction studies in patients referred with a clinical diagnosis of carpal tunnel syndrome. *J Clin Ultrasound.* (2009) 37:389–93. doi: 10.1002/jcu.20601
24. French C, Cartwright MS, Hobson-Webb LD, Boon AJ, Alter KE, Hunt CH, et al. Evidence-based guideline: neuromuscular ultrasound for the diagnosis of carpal tunnel syndrome. *Muscle Nerve.* (2012) 46:287–93. doi: 10.1002/mus.23389
25. Kwon BC, Jung KI, Baek GH. Comparison of sonography and electrodiagnostic testing in the diagnosis of carpal tunnel syndrome. *J Hand Surg Am.* (2008) 33:65–71. doi: 10.1016/j.jhsa.2007.10.014
26. Charles K. Nerve conduction studied and electromyograph for the evaluation of patients with carpal tynnel syndrome. *Muscle Nerve.* (1993) 16:1392–414. doi: 10.1002/mus.880161220
27. Claes F, Kasius KM, Meulstee J, Grotenhuis JA, Verhagen WIM. Treatment outcome in carpal tunnel syndrome: does distribution of sensory symptoms matter? *J Neurol Sci.* (2014) 344:143–8. doi: 10.1016/j.jns.2014.06.044
28. Vanderschueren GAKB, Meys VEWH, Beekman R. Doppler sonography for the diagnosis of carpal tunnel syndrome: a critical review. *Muscle Nerve.* (2014) 50:159–63. doi: 10.1002/mus.24241
29. Roomizadeh P, Eftekharsadat B, Abedini A, Ranjbar-Kiyakalayeh S, Yousefi N, Ebadi S, et al. Ultrasonographic assessment of carpal tunnel syndrome severity: a systematic review and meta-analysis. *Am J Phys Med Rehabil.* (2019) 98:373–81. doi: 10.1097/PHM.0000000000001104
30. Hobson-Webb LD, Massey JM, Juel VC, Sanders DB. The ultrasonographic wrist-to-forearm median nerve area ratio in carpal tunnel syndrome. *Clin Neurophysiol.* (2008) 119:1353–7. doi: 10.1016/j.clinph.2008.01.101
31. Visser LH, Smidt MH, Lee ML. Diagnostic value of wrist median nerve cross sectional area versus wrist-to-forearm ratio in carpal tunnel syndrome. *Clin Neurophysiol.* (2008) 119:2898–9. doi: 10.1016/j.clinph.2008.08.022
32. Chang YW, Hsieh TC, Tzeng IS, Chiu V, Huang PJ, Horng YS. Ratio and difference of the cross-sectional area of median nerve to ulnar nerve in diagnosing carpal tunnel syndrome: a case control study. *BMC Med Imaging.* (2019) 19:52. doi: 10.1186/s12880-019-0351-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 16 December 2020
doi: 10.3389/fneur.2020.613446

Intensive Teenage Activity Is Associated With Greater Muscle Hyperintensity on T1W Magnetic Resonance Imaging in Adults With Dysferlinopathy

*Ursula Moore[1], Marni Jacobs[2,3], Roberto Fernandez-Torron[1,4], Jaume LLauger Rossello[5], Fiona E. Smith[6], Meredith James[1], Anna Mayhew[1], Laura Rufibach[7], Pierre G. Carlier[8], Andrew M. Blamire[6], John W. Day[9], Kristi J. Jones[10], Diana X. Bharucha-Goebel[11,12], Emmanuelle Salort-Campana[13], Alan Pestronk[14], Maggie C. Walter[15], Carmen Paradas[16,17], Tanya Stojkovic[18], Madoka Mori-Yoshimura[19], Elena Bravver[20†], Elena Pegoraro[21], Jerry R. Mendell[22], on behalf of The Jain COS Consortium, Kate Bushby[1], Volker Straub[1] and Jordi Diaz-Manera[1,23,24]**

[1] The John Walton Muscular Dystrophy Research Centre, Translational and Clinical Research Institute, Newcastle University, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom, [2] Division of Biostatistics and Study Methodology, Center for Translational Science, Children's National Health System, Washington, DC, United States, [3] Pediatrics, Epidemiology and Biostatistics, George Washington University, Washington, DC, United States, [4] Neuromuscular Area, Biodonostia Health Research Institute, Neurology Service, Donostia University Hospital, Donostia-San Sebastian, Spain, [5] Radiology Department, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain, [6] Magnetic Resonance Centre, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom, [7] Jain Foundation, Newcastle upon Tyne, Seattle, WA, United States, [8] Université Paris-Saclay, CEA, DRF, Service Hospitalier Frederic Joliot, Orsay, France, [9] Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, United States, [10] The Children's Hospital at Westmead, The University of Sydney, Sydney, NSW, Australia, [11] Department of Neurology Children's National Health System, Washington, DC, United States, [12] National Institutes of Health (NINDS), Bethesda, MD, United States, [13] Service des maladies neuromusculaire et de la SLA, Hôpital de La Timone, Marseille, France, [14] Department of Neurology Washington University School of Medicine, St. Louis, MO, United States, [15] Department of Neurology, Friedrich-Baur-Institute, Ludwig-Maximilians-University of Munich, Munich, Germany, [16] Center for Biomedical Network Research on Eurodegenerative Diseases, Instituto de Salud Carlos III, Madrid, Spain, [17] Neuromuscular Unit, Department of Neurology, Hospital U. Virgen del Rocío/Instituto de Biomedicina de Sevilla, Sevilla, Spain, [18] Centre de référence des maladies neuromusculaires, Institut de Myologie, AP-HP, Sorbonne Université, Hôpital Pitié-Salpêtrière, Paris, France, [19] Department of Neurology, National Center Hospital, National Center of Neurology and Psychiatry Tokyo, Tokyo, Japan, [20] Carolinas Neuromuscular/ALS-MDA Center, Neuroscience Institute, Carolinas HealthCare System, Charlotte, NC, United States, [21] Department of Neuroscience, University of Padova, Padova, Italy, [22] The Abigail Wexner Research Institute at Nationwide Children's Hospital, Columbus, OH, United States, [23] Neuromuscular disorders Unit, Department of Neurology, Hospital de la Santa Creu I Sant Pau, Barcelona, Spain, [24] Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER), Valencia, Spain

Practice of sports during childhood or adolescence correlates with an earlier onset and more rapidly progressing phenotype in dysferlinopathies. To determine if this correlation relates to greater muscle pathology that persists into adulthood, we investigated the effect of exercise on the degree of muscle fatty replacement measured using muscle MRI. We reviewed pelvic, thigh and leg T1W MRI scans from 160 patients with genetically confirmed dysferlinopathy from the Jain Foundation International clinical outcomes study in dysferlinopathy. Two independent assessors used the Lamminen-Mercuri visual scale to score degree of fat replacement in each muscle. Exercise intensity for each individual was defined as no activity, minimal, moderate, or intensive activity by using metabolic equivalents and patient reported frequency of sports undertaken between

OPEN ACCESS

Edited by:
Xin-Ming Shen,
Mayo Clinic, United States

Reviewed by:
Anna Pichiecchio,
Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy
Teerin Liewluck,
Mayo Clinic, United States

***Correspondence:**
Jordi Diaz-Manera
jordi.diaz-manera@newcastle.ac.uk

[†] deceased

Specialty section:
This article was submitted to Neuromuscular Diseases, a section of the journal Frontiers in Neurology

Received: 02 October 2020
Accepted: 23 November 2020
Published: 16 December 2020

Citation:
Moore U, Jacobs M, Fernandez-Torron R, LLauger Rossello J, Smith FE, James M, Mayhew A, Rufibach L, Carlier PG, Blamire AM, Day JW, Jones KJ, Bharucha-Goebel DX, Salort-Campana E, Pestronk A, Walter MC, Paradas C, Stojkovic T, Mori-Yoshimura M, Bravver E, Pegoraro E, Mendell JR, Bushby K, Straub V and Diaz-Manera J (2020) Intensive Teenage Activity Is Associated With Greater Muscle Hyperintensity on T1W Magnetic Resonance Imaging in Adults With Dysferlinopathy. Front. Neurol. 11:613446. doi: 10.3389/fneur.2020.613446

the ages of 10 and 18. We used ANCOVA and linear modeling to compare the mean Lamminen-Mercuri score for the pelvis, thigh, and leg between exercise groups, controlling for age at assessment and symptom duration. Intensive exercisers showed greater fatty replacement in the muscles of the pelvis than moderate exercisers, but no significant differences of the thigh or leg. Within the pelvis, Psoas was the muscle most strongly associated with this exercise effect. In patients with a short symptom duration of <15 years there was a trend toward greater fatty replacement in the muscles of the thigh. These findings define key muscles involved in the exercise-phenotype effect that has previously been observed only clinically in dysferlinopathy and support recommendations that pre-symptomatic patients should avoid very intensive exercise.

Keywords: dysferlinopathy, Magnetic Resonace Imaging (MRI), Miyoshi myopathy, LGMDR2, limb girdle muscle dystrophy, exercise

INTRODUCTION

Dysferlinopathy is an autosomal recessively inherited form of muscular dystrophy caused by mutations in the *DYSF* gene. It usually presents in early adulthood and is characterized by progressive weakness and wasting of skeletal muscles (1).

Unlike many other forms of muscular dystrophy, patients with dysferlinopathy often report performing a large amount of physical activity as children and young adults, before their symptoms first present (2, 3). This intensive exercise appears to be detrimental and we have previously demonstrated that higher levels of exercise before symptom onset is associated with both earlier patient reported symptom onset and earlier subsequent wheelchair requirement (4).

The association of exercise and disease progression in dysferlinopathy has been investigated in mouse models, with interesting results. Eccentric muscle contractions (such as running) cause more rapid progression of muscular dystrophy pathology and greater functional weakness, while concentric muscle contraction (such as swimming) appeared to be protective against both general disease progression and in mitigating the myofiber damage caused by subsequent eccentric muscle contraction (5, 6). This finding clearly has implications for patients in terms of the type or intensity of exercise that should be recommended. However, murine models of dysferlinopathy are not particularly good mimics of the human phenotype, being much less severe, and it is not certain if these findings can be generalized (6). In humans with dysferlinopathy, pathological differences depending on exercise type or intensity have not been assessed and it is not clear if early influences of exercise would remain detectable after many years of symptomatic disease.

Magnetic resonance imaging (MRI) can be used to non-invasively and objectively measure increases in intramuscular fat, acting as marker for muscle pathology. T1W sequences clearly delineate fat and muscle and are used to produce images for visual inspection to determine the degree of the muscle that has been replaced by fat. This method is used to determine both the pattern and severity of muscle involvement in neuromuscular disorders, and is widely used to support genetic diagnosis (7, 8). T1W MRI has been used to characterize a distinctive pattern of muscle involvement in dysferlinopathy, with certain muscles consistently demonstrating greater, or earlier, fat replacement than others (9).

In this study we use T1W images from this previously described cohort, in combination with questionnaire derived information about pre-symptomatic activity levels to objectively assess the impact of teenage exercise on muscle fat replacement. We review if this supports the patient reported earlier onset and more rapid loss of ambulation in intensive exercisers, if exercise has differential effects in muscle, such that some muscles are more impacted than others and if conclusions can be reached about the types of exercise that may be detrimental in dysferlinopathy.

METHODS

Study Subjects

This investigation reviews MRI imaging from the Jain Foundation's international longitudinal *Clinical Outcomes Study for Dysferlinopathy* (COS). This study included 201 patients with genetically confirmed dysferlinopathy from 15 sites internationally (10). This study received ethical approval in all participating countries. MRI was not mandatory for inclusion in the study and some sites could only perform lower limb MRI (not including the pelvis). Overall, 182 patients had a baseline MRI, including 84 patients with a whole body scan and 98 patients with only lower limb scans. Imaging from 22 patients did not produce adequate images for visual scoring to be applied (see below) and were excluded. This investigation therefore reviews MRI imaging from 160 patients.

This report uses the questionnaire based exercise information (**Supplementary Material**) and the MRI images collected from the patients screening and baseline visits, respectively. These visits also involved medical interview and examination, physiotherapy assessment, blood sampling, cardiac, and respiratory investigations. A cross-sectional description of the baseline cohort demographic and functional information (10), and pattern of muscle involvement on T1W MRI at baseline (9) have been previously published.

Semiquantitative MRI Assessment

Semiquantitative assessments of MRIs were performed by a blinded neurologist (RF-T) and radiologist (JL), who independently evaluated axial T1-weighted sequences with the Lamminen-Mercuri visual scale, with an inter-rater agreement kappa of 0.93 (95% CI 0.91–0.96), as previously reported (9). For the 4% of scans where observers scores did not match, observers reviewed the muscles together and agreed a final score (9). All 160 patients had complete Lamminen-Mercuri score results for the lower leg muscles—Tibialis anterior, extensor digitorum, peroneus longus, peroneus brevis, gastrocnemius medialis and lateralis, soleus, flexor digitorum, and tibialis posterior. Complete imaging of adductors was not included in some patients, leaving 106 of these 160 patients who also had a complete set of Lamminen-Mercuri scores for thigh muscles (Adductor brevis, longus an major, quadratus femoris, rectus femoris, vastus lateralis, vastus intermedius and vastus medialis, sartorius, gracilis, semimembranosus, semitendinosus, long, and short head of biceps femoris). Full pelvic muscle Lamminen-Mercuri scores were available for 67 of the 160. Pelvic muscles were psoas, pectineus, piriformis, iliacus, gluteus medius, gluteus maximus, gluteus minor, tensor fasciae latae, obturator internus, and obturator externus.

Exercise Scoring

At the screening visit, patients reported the type and frequency of all regular activities performed as children and teenagers (**Supplementary Material**). Exercise was coded based on the maximum patient reported frequency between the age of 10 and 18 years, and the metabolic equivalent (MET) of the activities described, as previously reported (4). Group 0 reported no physical activity, group 1 reported vigorous activity occasionally/monthly, or moderate activity once a week, group 2 reported moderate activity multiple times a week or vigorous activity once weekly, and group 3 reported vigorous activity multiple times per week.

Statistical Analysis

Age at MRI and symptom duration were not normally distributed and so were compared in a stepwise fashion between exercise groups using the Wilcoxon-Mann-Whitney test for non-parametric unpaired samples (i.e., group 0 vs. group 1, group 0 vs. group 2 and so forth for comparison between all groups).

A mean Lamminen-Mercuri score for the distal leg was calculated from the sum of scores of each individual muscle ($n = 160$). This was repeated to generate a mean Lamminen-Mercuri score for the thigh ($n = 106$) and the pelvis ($n = 67$). Muscles included are listed above.

To determine if there was any difference in Lamminen-Mercuri scores between exercise groups, mean Lamminen-Mercuri scores of the a. leg, b. thigh and c. pelvis were compared between exercise groups using type III ANCOVA, using the "Anova" function from the package "car" in the programme R. Age at MRI and symptom duration were assessed as covariates.

TABLE 1 | Table showing demographic factors of patients in each exercise group at baseline.

	Exercise group			
	0	**1**	**2**	**3**
Distal leg MRI ($n = 160$)				
Number of patients	39	17	54	50
Male (%)	14 (36)	9 (53)	30 (56)	28 (56)
Symptom onset age in years—median (range)*	25.00 (12–48)	17.00 (12–60)	18.00 (10–41)	18.00 (12.5–40)
Symptom duration in years—median (range)**	20.00 (4–41)	11.00 (2–42)	12.50 (2–38)	17.00 (3–51)
Age at assessment in years—median (range)***	47.00 (22–86)	35.00 (15–71)	31.50 (15–57)	37.50 (22–67)
Thigh MRI in ($n = 106$)				
Number of patients	25	12	40	29
Male (%)	9 (36)	6 (50)	23 (58)	17 (59)
Symptom onset age in years – median (range)(*)	25.00 (12–48)	18.5 (12–60)	18.00 (10–41)	17.5 (13–28)
Symptom duration in years—median (range)	21(4–41)	18.5 (3–42)	13.00 (2–38)	16.5 (3–51)
Age at assessment in years—median (range)****	47.00 (32–86)	37.5 (15–71)	30.00 (15–57)	37.00 (22–64)
Pelvic MRI ($n = 67$)				
Number of patients	11	6	28	22
Male (%)	2 (18)	3 (50)	13 (46)	12 (55)
Symptom onset age in years—median (range)*	26.00 (14–39)	16.00 (12–60)	18.00 (12–41)	17.50 (13–27)
Symptom duration in years—median (range)**	16.00 (10–41)	14.00 (3–26)	13.00 (4–23)	19.00 (6–51)
Age at assessment in years—median (range)***	47.00 (32–66)	32.50 (15–71)	32.500 (19–57)	39.00 (22–64)

MRI: T1 weighted magnetic resonance imaging.

**Median age of symptom onset significantly higher in group 0 compared to group 1, 2, and 3.*

***Median disease duration significantly shorter in group 2 than in group 0 or group 3.*

****Group 0 are significantly older than group 3 who are, in turn, significantly older than group 2.*

*****Group 0 is significantly older than group 2.*

Firstly, age was considered as a covariate to determine if patients in a particular exercise group at any given age had a different Lamminen-Mercuri score than those in another group. Secondly, symptom duration and age were combined as covariates to determine if fat replacement on MRI progressed more rapidly (from the onset of symptoms) in one exercise group than another. *Post-hoc* analysis of significant ($p < 0.05$) ANCOVA results was completed using the "lsmeans()" function, from the package "lsmeans." This was to identify which exercise groups were significantly different from each other, and thus causing the significant ANCOVA result.

For visual representation of the exercise dependent effects, linear modeling was used. A linear model of Lamminen-Mercuri score was generated with age and exercise group as covariates. This allowed calculation of the mean model predicted Lamminen-Mercuri score for each exercise group as if all patients were of the same age (39 years—the mean age of the cohort). This was repeated with disease duration as an additional covariate (along with age) to produce model predicted Lamminen-Mercuri scores as if all patients had had the same disease duration (17 years—the mean disease duration of the cohort) and age.

Muscle groups showing significant differences between exercise groups were then further reviewed with additional ANCOVA of the Lamminen-Mercuri scores of the individual muscles within that group and compared in the same way as for the muscle groups.

In order to review if results differed in a subset of patients with a shorter disease duration, we repeated this analysis of pelvic, thigh and leg muscle groups in patients with symptoms for <15 years at the time of assessment.

Post-hoc review of *p*-values using Bonferonni correction for multiple comparisons was performed. No power calculations were conducted to determine group size as data was from an existing study.

RESULTS

Demographics

Patient reported age of symptom onset was later in the inactive group 0 than in all of the other exercise groups (**Table 1**). Patients in group 0 were significantly older at the time of assessment than those in the extremely active group 3, who were in turn older than those in group 2 ($p < 0.05$) (**Table 1**). Overall, patients in the moderately active group 2 had a shorter median symptom duration at the time of assessment than those in group 0 or in group 3 ($p < 0.05$). Number of patients in group 1 was small (17 patients in distal leg MRI group) with variable age and disease

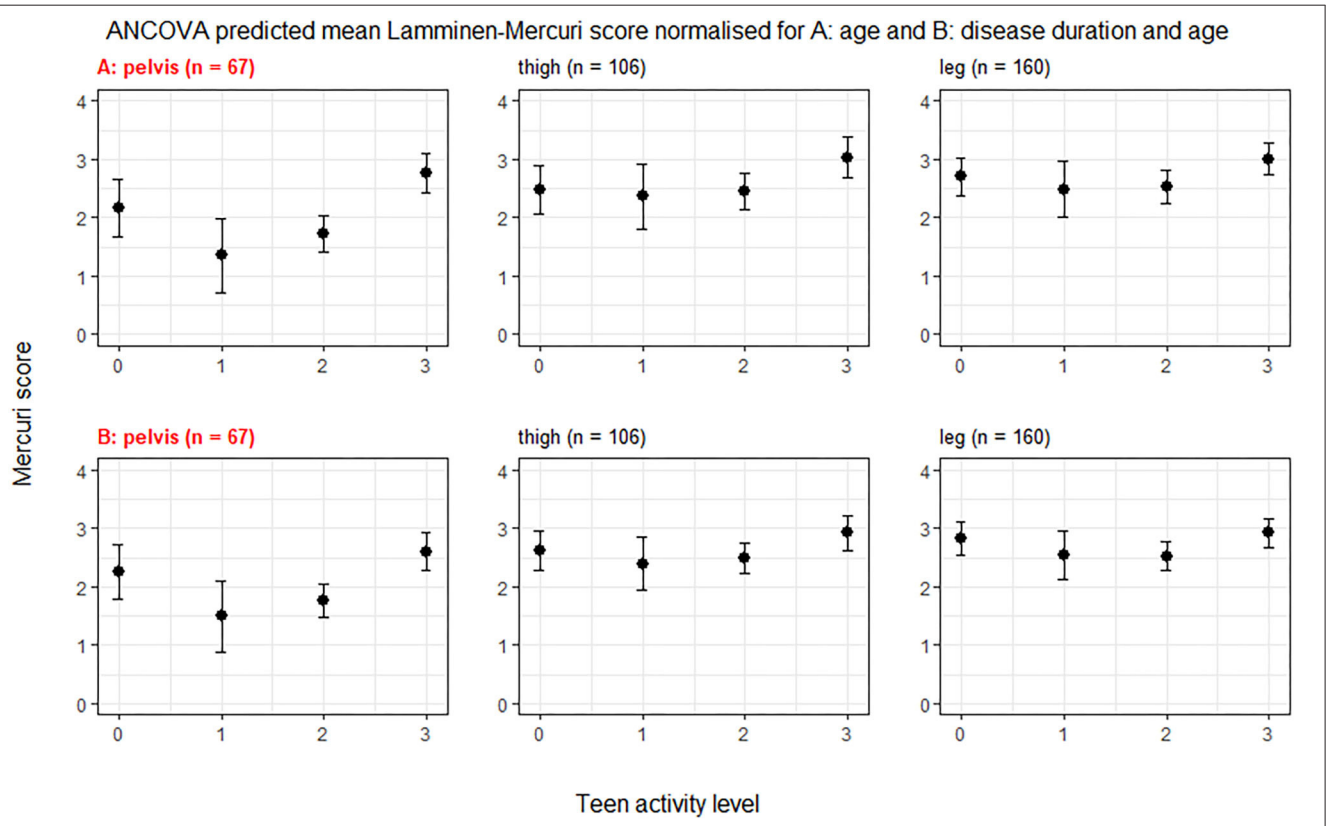


FIGURE 1 | ANCOVA predicted Lamminen-Mercuri scores for each exercise group. The top row **(A)** shows predicted score in each exercise group (0, 1, 2, or 3) adjusted for age—i.e., as though all patients were the cohort mean age of 39 years. The bottom row **(B)** shows predicted score in each exercise group (0, 1, 2, or 3) adjusted for disease duration and age—i.e., as though all patients had an age of 39 years and the cohort mean symptom duration of 17 years. Bars show the 95% confidence interval. Red text denotes groups showing significant differences on ANCOVA ($p < 0.005$).

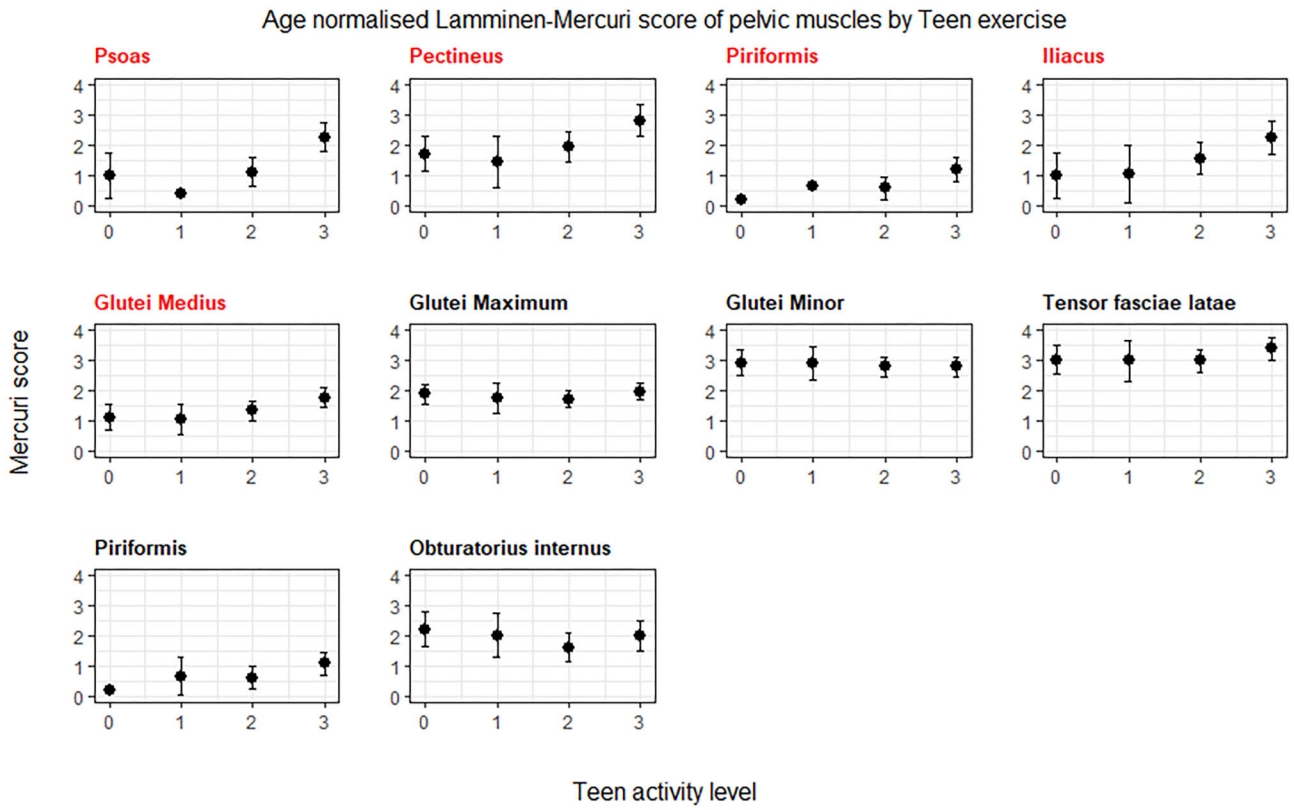


FIGURE 2 | ANCOVA predicted Lamminen-Mercuri scores of the pelvis muscles, adjusted for age. Bars show the 95% confidence interval. Red text denotes groups showing significant differences on ANCOVA ($p < 0.005$). Patients are split into exercise groups 0, 1, 2, or 3.

duration, which was not significantly different from those in other exercise groups.

In the groups of patients who had pelvic or thigh imaging, the median age of onset, age at assessment and symptom duration in each exercise group did not differ significantly from that in the larger cohort.

Age Adjusted Mercuri Score

There were significant differences between exercise groups in the mean predicted Lamminen-Mercuri score of the pelvic muscles ($p \leq 0.0001$), when controlling for age as a covariate. This was driven by a significantly higher predicted Lamminen-Mercuri score in exercise group 3, compared to group 1 (3 vs. 1, $p = 0.0018$) and group 2 (3 vs. 2, $p = 0.0002$). There was no significant difference between group 0 and any of the other exercise groups. There was no difference in mean predicted Lamminen-Mercuri score of thigh or leg muscles between groups (**Figure 1**).

Modeling using age and teenage exercise level could account for one third of the observed variability in Lamminen-Mercuri score in the whole leg (adjusted $R^2 = 0.33$), while using age alone produced an R^2 value of 0.14.

The individual muscles in the pelvis showing a significant difference ($p < 0.05$) in age normalized Lamminen-Mercuri score between exercise groups were psoas, piriformis, pectineus, iliacus, and gluteus medius (**Figure 2**). However, after correction for multiple comparisons, this only remained significant for psoas ($p = 0.0004$), which showed a significantly higher predicted Lamminen-Mercuri score in exercise group 3 compared to groups 0 ($p = 0.03$), 1 ($p = 0.0032$), and 2 ($p = 0.007$).

In the smaller subset of patients ($n = 73$) with symptoms for <15 years, there was a trend toward greater Lamminen-Mercuri score in groups 3 than 2 and in group 2 than 1 in the pelvis and thigh. However, there was no longer a significant difference in pelvic muscles by exercise group and although there was a difference between groups in the thigh muscles ($p = 0.044$), this did not remain after correction for multiple comparisons (**Figure 3**).

Symptom Duration and Age Adjusted Mercuri Score

In ANCOVA of Lamminen-Mercuri score with symptom duration, age and exercise group as covariates, symptom duration was always a stronger predictor of Lamminen-Mercuri score than age and age was not an independent predictor of Lamminen-Mercuri score.

Symptom duration was a strong predictor of Lamminen-Mercuri score and modeling using symptom duration alone accounted for 38% of the variation in Lamminen-Mercuri score (adjusted $R^2 = 0.38$) in the whole leg. Adding exercise group to the model improved the adjusted R^2 value to 0.47, which was not

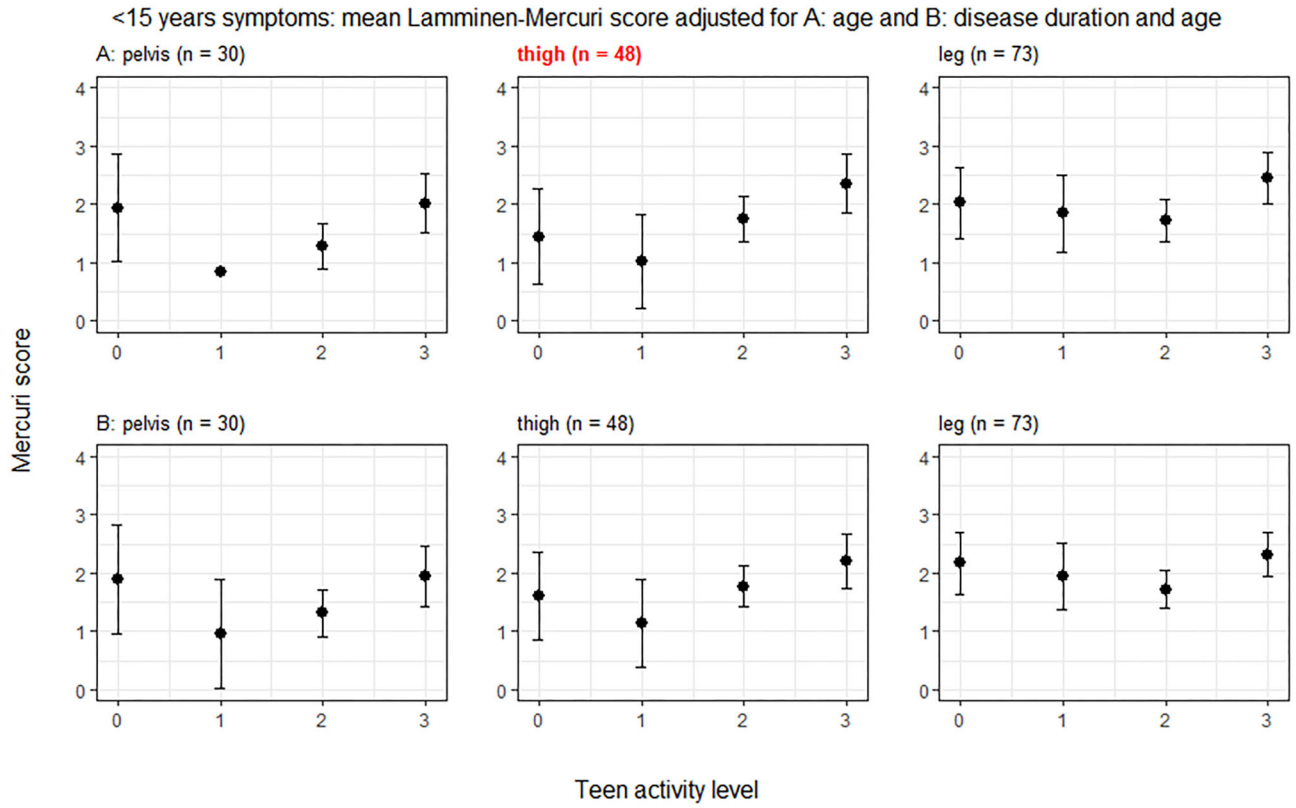


FIGURE 3 | ANCOVA predicted Lamminen-Mercuri scores for each exercise group in those with <15 years of symptoms. The top row **(A)** shows predicted score in each exercise group (0, 1, 2, or 3) adjusted for age—i.e., as though all patients were the cohort mean age of 31 years. The bottom row **(B)** shows predicted score in each exercise group (0, 1, 2, or 3) adjusted for disease duration and age—i.e., as though all patients had an age of 31 years and the cohort mean symptom duration of 8 years. Bars show the 95% confidence interval. Red text denotes groups showing significant differences on ANCOVA ($p < 0.005$).

bettered by the further addition of age to the model (adjusted $R^2 = 0.47$). When accounting for symptom duration and age, there remained significant differences in mean predicted Lamminen-Mercuri score of the pelvic muscles ($p = 0.0015$) between exercise groups. Again, this was driven by a significantly higher predicted Lamminen-Mercuri score in exercise group 3, compared to group 2 (3 vs. 2, $p = 0.0029$) and group 1 (3 vs. 1, $p = 0.0148$). There was no significant difference between group 0 and any of the other exercise groups in the pelvis. There were no differences between exercise groups in thigh or leg muscle groups (**Figure 1**).

The muscle in the pelvis showing the greatest difference in age and symptom duration adjusted Lamminen-Mercuri score between exercise groups remained psoas ($p = 0.016$) (**Figure 4**, example MRI images in **Figure 5**). However, differences between groups in psoas Lamminen-Mercuri score were not significant on ANCOVA after correction for multiple comparisons.

In the smaller subset of patients with symptoms for <15 years, there was no longer a significant difference between exercise groups in any of the muscle groups.

DISCUSSION

We have demonstrated an association between intensive teenage exercise and a subsequent more severely affected MRI appearance of pelvic musculature in adult patients with dysferlinopathy, long after their teenage years. Psoas, pectineus, and piriformis were the most differentially affected muscles. Patients who performed intensive exercise as teenagers generally demonstrated a higher (worse) Lamminen-Mercuri score at a given age and disease duration, suggesting a more rapid progression of fat replacement from symptom onset.

We did not demonstrate any association between exercise group and the muscle groups of the thigh or distal leg in the overall group, but there was a trend toward a higher Lamminen-Mercuri score in the thigh muscle of intensive exercisers with <15 years symptom duration. The thigh and leg muscles generally have a higher Lamminen-Mercuri score in this cohort than the muscles of pelvis (9), suggesting earlier or more significant involvement of these muscles in the disease process. It may be that no difference was demonstrated between exercise groups in the thigh and distal leg overall because too many patients are already at a stage where these muscles are severely affected, essentially reaching a ceiling of the maximum Lamminen-Mercuri score of 4 relatively early in the disease process. The finding of potential differences in the thigh muscles in less advanced disease suggests that there may have been differences in all muscles at some point in progression of the disease.

In earlier work, we showed that symptom onset and age of part and full time wheelchair use is significantly earlier in intensive

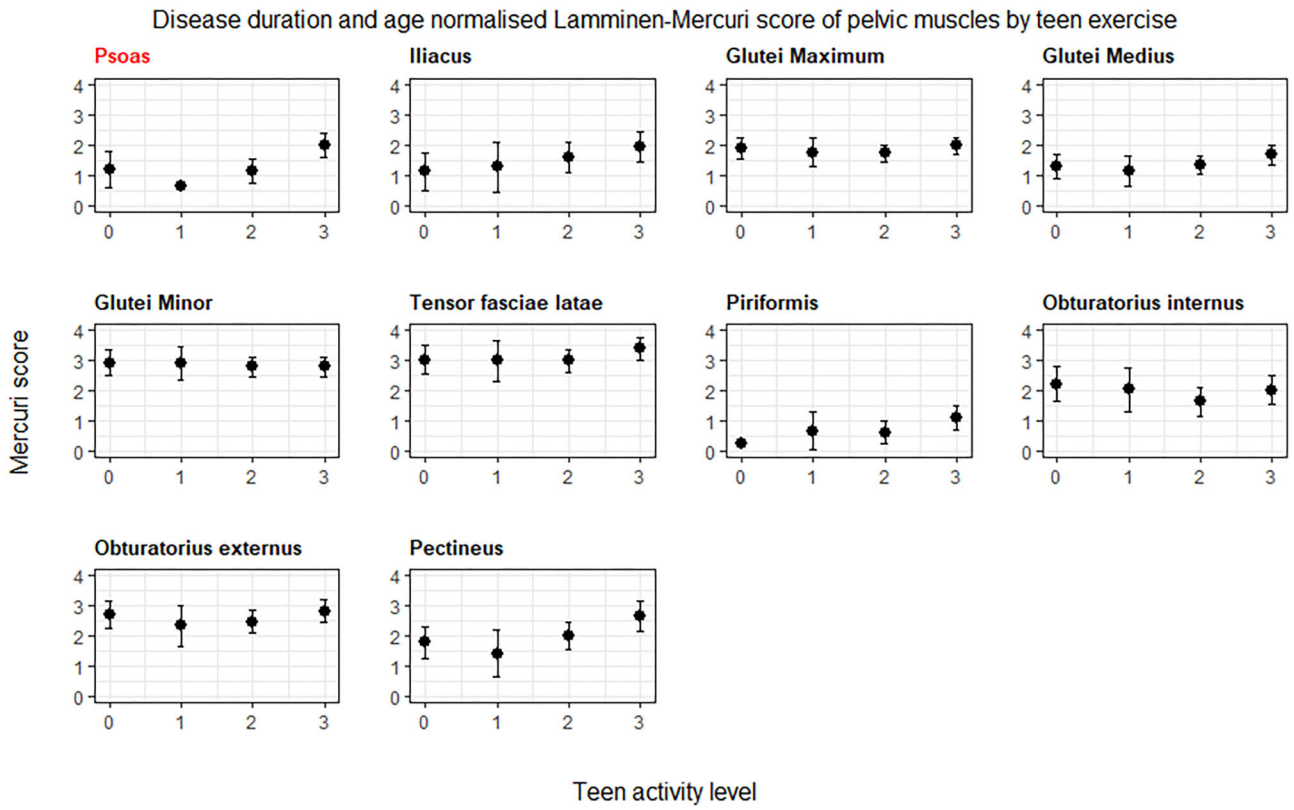


FIGURE 4 | ANCOVA predicted Lamminen-Mercuri scores of the pelvis muscles, adjusted for age, and disease duration. Bars show the 95% confidence interval. Red text denotes groups showing significant differences on ANCOVA ($p < 0.005$). Patients are split into exercise groups 0, 1, 2, or 3.

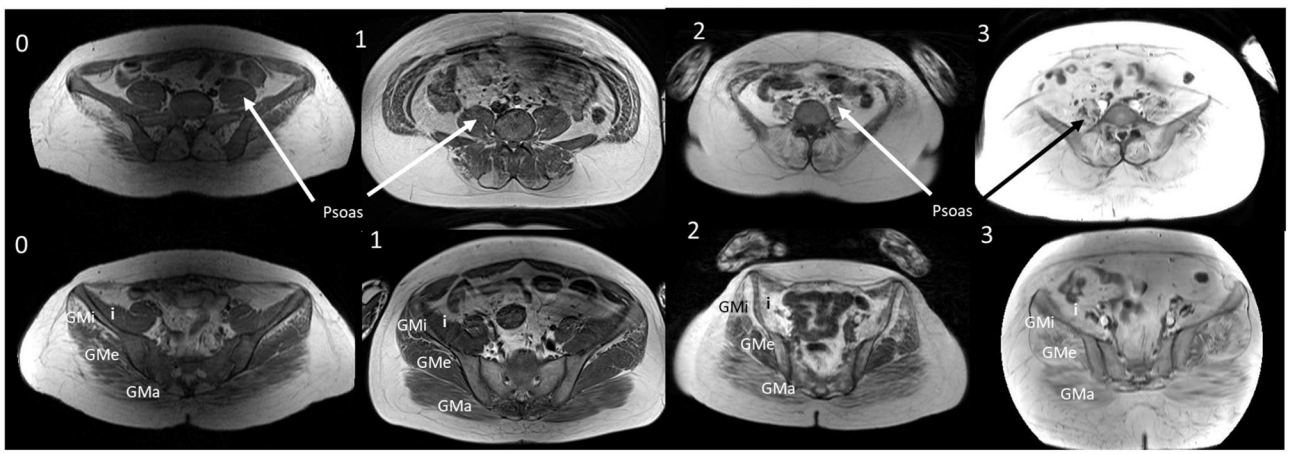


FIGURE 5 | T1W MRI images from age and duration matched patients. T1W MRI images of the lower abdomen and pelvisfor a representative patient from each exercise group (0, 1, 2, and 3) of a similar age and symptom duration. GMa, gluteus maximus; GMe, gluteus medius; GMi, gluteus minimus; I, Iliacus. Representative patients from groups 0, 2, and 3 were all aged 36 years with symptom duration 18 or 19 years at the time of the MRI. Group 1 patient was aged 33 years with a symptom duration of 21 years.

than moderate exercisers (4). While this MRI study adds objective evidence to this observation, we did not observe the previously described "dose" type of association with exercise—with non and very light exercisers (groups 0 and 1) having a significantly later age of onset and wheelchair requirement than those in group 2. We had anticipated that when investigating the MRI, this may translate to differences in Lamminen-Mercuri score between exercise group 0, group 2, and group 3. However, this was not seen and the key differences in Lamminen-Mercuri score were between exercise groups 2 and 3 and not group 0. There are several possible explanations for this. Firstly, this may be related to the possible ceiling effect of the Lamminen-Mercuri

score, as patients in group 0 were significantly older and had had symptoms for significantly longer than patients in group 2 and 3, making them more likely to have reached this plateau. It is also possible that, the more time that intervenes between teenage years and subsequent assessment the less marked the effect of exercise on imaging results—this would again dilute any effect in the older group 0. Finally, while the majority of non-exercisers had a relatively late onset of symptoms, some had a younger onset and probably did not exercise precisely because they had already started to experience symptoms. The muscle MRI of these more severely affected patients may "skew" the results of the non-exercising group 0.

The method used here to quantify the "amount" of exercise performed used the relative metabolic equivalents of the sports performed, giving an estimate of the intensity of the exercise. However, research in mice suggests that it may be the type of muscle contraction, rather than the metabolic cost, which determines the effect on dysferlin deficient muscle (5, 6). Unfortunately we were not able to assess this here because, as anticipated in an observational study, none of our participants had performed exclusively concentric exercise (such as swimming) without also reporting frequent eccentric sports (such as running).

We have identified specific muscles in the pelvis that were more affected in intensive exercisers in this cohort. This raises the question of why these particular muscles are affected while others are not. This may in part be due to differences only being observable in less affected muscle, as discussed above. However, there were some muscles which are not extensively involved across exercise groups (such as gluteus maximum) and yet still do not demonstrate differential involvement. It seems reasonable that sports that preferentially use specific muscles could affect these muscles more than others, and as many intensive sports involve a significant amount of running, intensive exercisers may see proportionally greater damage in muscles involved heavily in running—such as psoas (11, 12). We did attempt to review this in our cohort, comparing Lamminen-Mercuri scores for individual muscles between patients who reported different activities (e.g., predominantly running vs. predominantly swimming). We did not find differential Lamminen-Mercuri scores between the most frequently performed activities, however this analysis is likely significantly confounded by the multiple activities performed by each patient.

Muscle MRI offers the possibility to analyse the impact of exercise in muscle structure. In our opinion, the results of this study should encourage further research in muscle MRI biomarkers in dysferlinopathy patients performing different types of exercise. Muscle MRI sequences that could be helpful include but are not limited to STIR, T2 imaging, sodium MRI, and P31-spectroscopy (13–15). All of these previous imaging sequences have identified early changes in muscle structure, that in the case of dysferlinopathy patients, could help to elucidate which program of exercise is less harmful for the skeletal muscles.

This analysis confirms that T1W imaging demonstrates more severe fat replacement in muscles of patients with dysferlinopathy who performed intensive, rather than moderate, exercise as teenagers. This adds pathological evidence to the previous report linking symptom onset and earlier wheelchair use to intensive teenage exercise. For patients who receive a diagnosis before symptom onset, such as through family screening or detection of high CK, this evidence would support a recommendation to avoid very intensive exercise regimens before symptom onset, while maintaining a healthy lifestyle.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available on reasonable request to the study steering group. Requests should be directed to Professor Volker Straub at volker.straub@newcastle.ac.uk.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by local ethics committees in each country involved in the study: Comité Etico de investigacion con médicamentos de la Fundacio de Gestio Sanitaria de la Santa Creu I Sant Pau, Barcelona, Spain; The Carolinas HealthCare System Institutional Review Board Federal-Wide Assurance, Charlotte, United States; IRB at Nationwide Children's Hospital, Columbus, United States; Le Comité de Protection des Personnes Sud-Méditerranée, Marseille and Paris, France; Ethikkomission bei de LMU Munchen, Munich, Germany; Stanford University IRB, Panel on Human Subjects, Panel 7, Stanford, United States; The Washington University in St. Louis Institutional Review Board, St Louis, United States; The Medical Ethics Committee of the NCNP, Tokyo, Japan; Children's National Medical Center Institutional Review Board (Children's National IRB), Washing DC, United States; Newcastle and North Tyneside 2 medical ethics committee, Newcastle, United Kingdom; El comité de Etica de a investigacion de Centro H.U Virgen del Rocio de Sevilla, Seville, Spain; REGIONE VENETO AZIENDA OSPEDALIERA DI PADOVA Comitato Etico per la Sperimentazione, Padova, Italy; Children's Hospitals Network Human Research Ethics Committee, Sydney, Australia; North East—Newcastle and North Tyneside Health Research Authority research ethics committee, Newcastle, United Kingdom. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

UM contributed to data acquisition, conception and study design, analysis of data, and preparation of the manuscript. MaJ contributed to review of statistical analysis and drafting. RF-T, JL, FS, MeJ, AM, LR, PC, AB, JD, KJ, DB-G, ES-C, AP, MW, CP, TS, MM-Y, EB, EP, JM, and KB contributed to acquisition and analysis of data and drafting. VS contributed to conception and study design, acquisition and analysis of data, and drafting. JD-M contributed to conception and study design, acquisition and analysis of data, drafting, and final sign of corresponding author. All authors contributed to the article and approved the submitted version.

FUNDING

The estimated $4 million USD needed to fund this study was provided by the Jain Foundation. VS was supported by an MRC strategic award to establish an International Centre for Genomic Medicine in Neuromuscular Diseases (ICGNMD) MR/S005021/1.

ACKNOWLEDGMENTS

This study has only been possible thanks to the international collaboration of several specialized centres promoted by the Jain Foundation. The Jain COS consortium would like to thank the study participants and their families for their invaluable contribution and would also like to acknowledge the ongoing support the Jain Foundation provides in the development, management, and analysis of this Study. The Jain Foundation, based in Seattle, USA, is entirely focused on LGMD2B/dysferlinopathy/Miyoshi Myopathy. The foundation does not solicit funding from patients, but instead funds research and clinical studies worldwide with the goal of finding treatments for dysferlinopathy. Please visit www.jain-foundation.org for more information about the foundation and if you are a patient suffering from dysferlinopathy, please consider enrolling into their interactive dysferlinopathy registry that seeks to build a strong, engaged, and supportive community (patients@jain-foundation.org).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2020.613446/full#supplementary-material

REFERENCES

1. Aoki M. *Dysferlinopathy*. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews(®)*. Seattle, WA: University of Washington, Seattle (2004).
2. Klinge L, Aboumousa A, Eagle M, Hudson J, Sarkozy A, Vita G, et al. New aspects on patients affected by dysferlin deficient muscular dystrophy. *J Neurolo Neurosurg Psychiat.* (2010) 81:946–53. doi: 10.1136/jnnp.2009.178038
3. Angelini C, Peterle E, Gaiani A, Bortolussi L, Borsato C. Dysferlinopathy course and sportive activity: clues for possible treatment. *Acta Myol.* (2011) 30:127–32.
4. Moore UR, Jacobs M, Fernandez-Torron R, Jang J, James MK, Mayhew A, et al. Teenage exercise is associated with earlier symptom onset in dysferlinopathy: a retrospective cohort study. *J Neurol Neurosurg Psychiatry.* (2018) 89:1224–6. doi: 10.1136/jnnp-2017-317329
5. Biondi O, Villemeur M, Marchand A, Chretien F, Bourg N, Gherardi RK, et al. Dual effects of exercise in dysferlinopathy. *Am J Pathol.* (2013) 182:2298–309. doi: 10.1016/j.ajpath.2013.02.045
6. Begam M, Roche R, Hass JJ, Basel CA, Blackmer JM, Konja JT, et al. The effects of concentric and eccentric training in murine models of dysferlin-associated muscular dystrophy. *Muscle Nerve.* (2020) 62:393–403. doi: 10.1002/mus.26906
7. Warman Chardon J, Díaz-Manera J, Tasca G, Bönnemann CG, Gómez-Andrés D, Heerschap A, et al. MYO-MRI diagnostic protocols in genetic myopathies. *Neuromusc Disord.* (2019) 29:827–41. doi: 10.1016/j.nmd.2019.08.011
8. Leung DG. Advancements in magnetic resonance imaging-based biomarkers for muscular dystrophy. *Muscle Nerve.* (2019) 60:347–60. doi: 10.1002/mus.26497
9. Diaz-Manera J, Fernandez-Torron R, Llauger J, James MK, Mayhew A, Smith FE, et al. Muscle MRI in patients with dysferlinopathy: pattern recognition and implications for clinical trials. *J Neurol Neurosurg Psychiatry.* (2018) 89:1071–81. doi: 10.1136/jnnp-2017-317488
10. Harris E, Bladen CL, Mayhew A, James M, Bettinson K, Moore U, et al. The Clinical Outcome Study for dysferlinopathy: an international multicenter study. *Neurol Genet.* (2016) 2:e89. doi: 10.1212/NXG.0000000000000089
11. Anderson FC, Pandy MG. Individual muscle contributions to support in normal walking. *Gait Posture.* (2003) 17:159–69. doi: 10.1016/S0966-6362(02)00073-5
12. Dorn TW, Schache AG, Pandy MG. Muscular strategy shift in human running: dependence of running speed on hip and ankle muscle performance. *J Exp Biol.* (2012) 215:1944–56. doi: 10.1242/jeb.064527
13. Carlier PG, Azzabou N, de Sousa PL, Hicks A, Boisserie J-M, Amadon A, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging follow-up of adult Pompe patients. *J Inherited Metab Dis.* (2015) 38:565–72. doi: 10.1007/s10545-015-9825-9
14. Reyngoudt H, Marty B, Caldas de Almeida Araújo E, Baudin PY, Le Louër J, Boisserie JM, et al. Relationship between markers of disease activity and progression in skeletal muscle of GNE myopathy patients using quantitative nuclear magnetic resonance imaging and (31)P nuclear magnetic resonance spectroscopy. *Quant Imaging Med Surg.* (2020) 10:1450–64. doi: 10.21037/qims-20-39
15. Constantinides CD, Gillen JS, Boada FE, Pomper MG, Bottomley PA. Human skeletal muscle: sodium MR imaging and quantification-potential applications in exercise and disease. *Radiology.* (2000) 216:559–68. doi: 10.1148/radiology.216.2.r00jl46559

Conflict of Interest: UM, MJa, LR, AB, and AP reports the grant from the Jain Foundation. JD reports the grant from the Jain Foundation, personal fees from Biogen, Ionis, Avexis, Roche, Sarepta, Sanofi, Genzyme, Scholar Rock, Pfizer plus patents from Athena Diagnostics. DB-G reports membership of the Gene Therapy Network (Avexis). MW reports advisory board membership for Avexis, Biogen, Novartis, Roche, Santhera, Sarepta, PTC Therapeutics, Ultragenyx, Wave Sciences, plus personal fees from Novartis, Biogen, Ultragenyx, Santhera, PTC Therapeutics, Ask Bio, Audentes Therapeutics, Fulcrum Therapeutics, GIG Consul, Guidepoint Global, Novartis, PTC, Gruenthal Pharma. EP reports grants, personal fees, and non-financial support from Santhera, personal fees, and non-financial support from Sarepta, Personal fees, and non-financial support from PTC pharmaceuticals all outside this submitted work. VS reports the Jain Foundation grant and other grants and personal fees from Sarepta Therapeutics.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 05 January 2021
doi: 10.3389/fneur.2020.613489

Quantitative Muscle MRI as Outcome Measure in Patients With Becker Muscular Dystrophy—A 1-Year Follow-Up Study

Aisha M. Sheikh, Karen Rudolf, Nanna Witting and John Vissing*

Department of Neurology, Copenhagen Neuromuscular Center, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

Introduction: With the advent of emerging molecular therapies for muscular dystrophies, the need for knowledge about natural history course of such diseases is of utmost importance in the preparation for future trials. However, for Becker muscular dystrophy such knowledge is scarce.

Objective: In this 1-year follow-up study, we examined disease progression in Becker muscular dystrophy by monitoring changes in MRI-assessed muscle fat fraction (FF) in axial and lower limb muscles and quantitative muscle strength of axial muscles.

Methods and Materials: Sixteen patients with Becker muscular dystrophy were investigated by (1) muscle strength of the trunk using a Biodex dynamometer and (2) Dixon muscle MRI of paraspinal and lower limb muscles. Quantitative MRI data was analyzed in two parts: The first part consisted of all participants ($N = 16$). The second analysis assessed two separate groups comprising lesser affected participants ($N = 5$) and more severely affected patients ($n = 11$).

Results: Trunk extension and flexion strength remained stable from baseline to follow-up. MRI did not show any significant increase in muscle FF % from baseline to follow-up in all patients, except for multifidus at the spinal level T12 ($p = 0.01$). However, when we analyzed the two subgroups, according to disease severity, FF% increased in the lesser severely affected group at L4/L5 erector spinae ($p = 0.047$), sartorius ($p = 0.028$), gracilis ($p = 0.009$), tibialis anterior ($p = 0.047$), peroneals ($p = 0.028$), and gastrocnemius medialis ($p = 0.009$), while the severely affected group only increased significantly at T12 multifidus ($p = 0.028$) and T12 erector spinae ($p = 0.011$). No difference in muscle strength was observed in the two subgroups.

Conclusion: Our results add to the existing knowledge about the natural rate of disease progression in BMD. As quantitative MRI was able to identify changes where strength assessment was not, MRI could be a strong biomarker for change in BMD. However, our findings show that it is important to stratify patients with BMD according to phenotype for future clinical trials.

Keywords: Becker muscular dystrophy, quantitative muscle MRI, muscle strength, fat fraction, outcome measure

OPEN ACCESS

Edited by:
Massimiliano Filosto, University of Brescia, Italy

Reviewed by:
Anna Pichiecchio, Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy
Giorgio Tasca, Catholic University of the Sacred Heart, Italy

****Correspondence:***
Aisha M. Sheikh
aisha.munawar.sheikh@regionh.dk

Specialty section:
This article was submitted to Neuromuscular Diseases, a section of the journal Frontiers in Neurology

Received: *02 October 2020*
Accepted: *26 November 2020*
Published: *05 January 2021*

Citation:
Sheikh AM, Rudolf K, Witting N and Vissing J (2021) Quantitative Muscle MRI as Outcome Measure in Patients With Becker Muscular Dystrophy—A 1-Year Follow-Up Study. Front. Neurol. 11:613489. doi: 10.3389/fneur.2020.613489

INTRODUCTION

With the advent of emerging molecular therapies for muscular dystrophies, there is a need to know how the natural history of these diseases is, but for Becker muscular dystrophy such knowledge is scarce, especially with regard to axial muscles.

Previous imaging studies examining the lower limb muscles, using visual rating scales in studies of patients with BMD, have shown distinctive patterns of muscle involvement of the hamstrings, quadriceps, and gastrocnemius (1, 2). To quantify pathological changes and to monitor disease progression in myopathies, quantitative MRI (qMRI) technique is superior to visual rating scales to show change (3, 4). Quantitative strength assessment of the lower limb muscles and the muscles of the trunk using Biodex has been established to be a reliable method in previous studies (5, 6). With this 1-year follow-up study, we wish to gain knowledge on how axial muscle involvement progresses over time in BMD using quantitative muscle MRI.

METHODS AND MATERIALS

Study Design and Participants

This prospective longitudinal study was conducted from April 2018 to June 2020 at Copenhagen Neuromuscular Center, Rigshospitalet, Copenhagen, Denmark, in accordance with the declaration of Helsinki and was approved by the Danish National Committee on Health Research Ethics (approval number: H-16030358).

We recruited 16 participants with genetically verified Becker muscular dystrophy from our Neuromuscular Center (**Table 1**). All 16 participants had previously participated in a cross-sectional study. Thirteen participants were able to ambulate without walking aid, while three were dependent on wheelchair for most functions. All participants were evaluated at baseline and at follow-up for trunk muscle strength and qMRI of the paraspinal and lower limb muscles. Muscle strength measure and qMRI were completed on the same day.

Quantitative Muscle Strength Measure

Maximal voluntary isometric contraction of the trunk was acquired using Biodex (Biodex System 4 Pro, Biodex Medical Systems, Shirley, NY) with a Dual position back Extension/Flexion attachment (model number 830-450). The anterior iliac spine was aligned with the attachment's fixed axis of rotation, and back support was set at 100 degrees of hip angle. To minimize the influence of muscles from other parts of the body, chest and thighs were immobilized with Velcro straps, and the participants were asked to cross their arms in front of their chest during the test. Each participant was instructed to perform a maximal isometric trunk extension and trunk flexion. One contraction lasted 5 s. To ensure maximal contraction, two submaximal test trials were performed to familiarize the participants with the testing protocol for each position followed by three trials of maximal contractions (interchangeably between extension and flexion) with 30 s of rest in between each contraction. Standardized verbal encouragement was provided to each participant during testing.

TABLE 1 | Demographics of the 16 participating Becker muscular dystrophy patients.

Subject ID	Age at follow-up	Age at symptom onset	Mutation	Disease duration in years
BMD 1	46	28	c.676_678del; p.Lys226del	18
BMD 2	51	41	Del26	10
BMD 3	37	10	Del45-48	27
BMD 4	31	10	c.6912+1G>T	21
BMD 5	38	2	Del45-48	36
BMD 6	33	29	c.1602 G>A	4
BMD 7	28	24	Del45-47	4
BMD 8	34	5	Del45-48	29
BMD 9	19	1	c.5632C>T, p.(Gln1878*)	18
BMD 10	39	22	Del45-48	17
BMD 11	39	6	Del45-48	33
BMD 12	30	29	Del45-47	1
BMD 13	32	8	Del45-47	24
BMD 14	26	6	Del45-49	20
BMD 15	60	35	Del48	25
BMD 16	60	7	Del45-47	53

Table displays age at follow-up, age at symptom onset, gene mutation, and disease duration in years.

MRI Data Acquisition and Processing

Images were acquired on a Siemens 3.0 Tesla Magnetom Verio scanner (Erlangen, Germany) at the Department of Radiology at the National University Hospital, Rigshospitalet in Copenhagen, Denmark. Each participant was examined in head-first supine position. The MRI protocol composed of a three-plane localizer sequence followed by a T1-weighted sequence (field of view (FOV), 400–450 mm; slice thickness, 6.0 mm; distance factor, 20%; repetition time (TR)/echo time (TE), 19/650 ms), and a 2-point Dixon sequence (FOV, 400–450 mm; slice thickness, 3.5 mm; distance factor, 0%; TE/TR, 2.45 and 3.675/5.59 ms). Two body matrix coils and a peripheral leg coil were used for signal detection. Total scan time was approximately 40 min. Five cross-sectional slices were chosen for the investigation of disease progression (**Figure 1**). Dixon sequences were used to quantify fat fraction (FF) by defining a region of interest (ROI). Twenty-two ROIs were mapped bilaterally using Horos software v. 3.3.6 and consisted of the following muscles: Erector spinae (ES) was mapped at spinal level C6, Th12, and L4/L5 and multifidus (M) at spinal levels T12 and L4/L5. Abdominal muscles were not mapped but instead inspected visually due to motion artifact from respiration. Iliopsoas (P) was mapped at spinal level L4/L5, and muscles of the thigh were mapped at mid-thigh, corresponding to 50% of the length of femur: Rectus femoris (RF), vastus lateralis (VL), vastus medialis (VM), vastus intermedius (VI), Sartorius (SA), gracilis (GR), the adductor muscles (AM) (adductor magnus and adductor longus), and hamstring muscles: semimembranosus (SM), semitendinosus (ST), and biceps femoris (BF). At the widest section of the

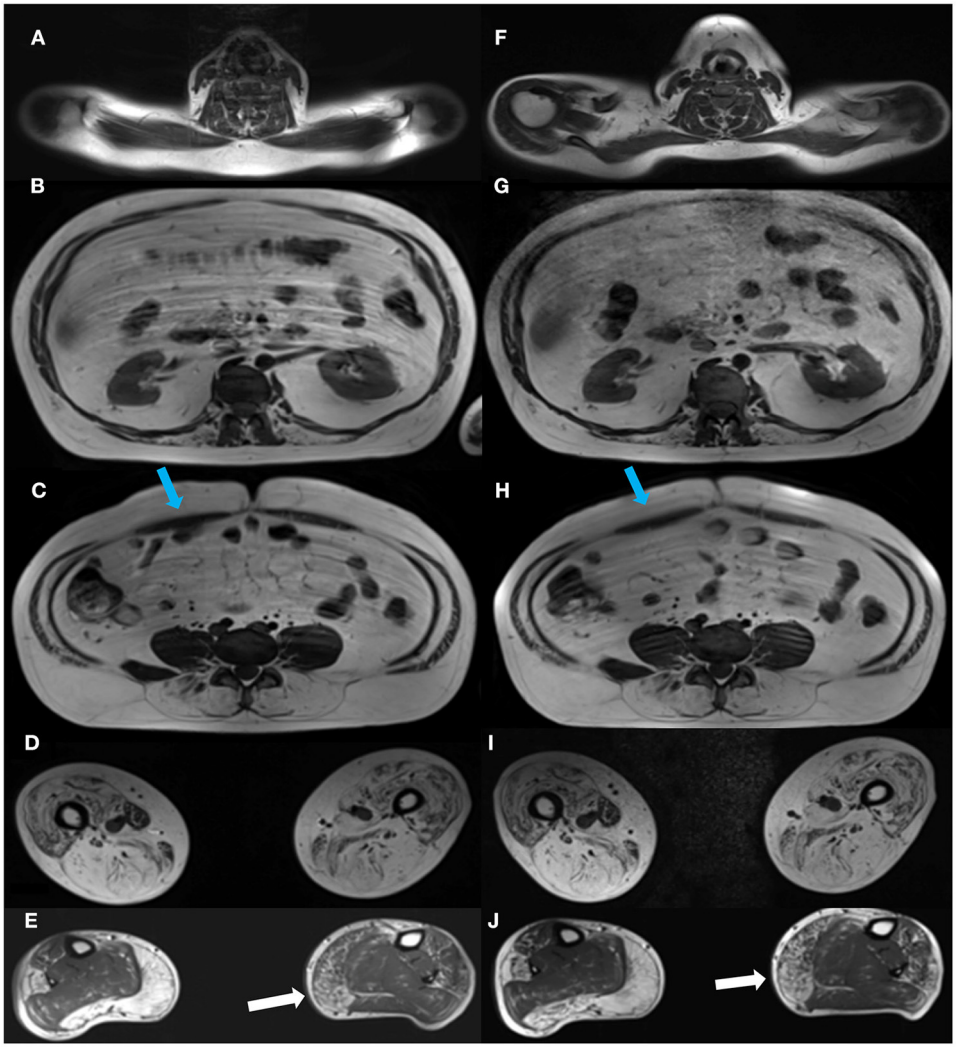


FIGURE 1 | Cross-sectional MR images at baseline and at follow-up. Images of the cross-sectional slices at baseline (left column, **A–E**) and at follow-up (right column, **F–J**). Images are shown at spinal levels C6 **(A,F)**, T12 **(B,G)**, and L4/L5 **(C,H)**, thighs **(D,I)**, and lower legs **(E,J)**. A visual inspection of the images shows some progression of the right medial head of gastrocnemius from baseline, **(E)** to follow-up **(J)**, while changes in other muscles from baseline are indiscernible. Blue arrows show the abdominal muscles **(C,H)**.

lower leg, corresponding to about 1/3 of tibia from the knee down, the following muscles were mapped: Tibialis anterior (TA), peroneals (PER), tibialis posterior (TP), soleus (SOL), gastrocnemius lateralis (GL), and gastrocnemius medialis (GM).

The cross-sectional area was determined from the mapping followed by quantitative fat fraction estimation. Mean FF % was expressed as percentage fat (0–100%) = signal fat/(signal water + fat). Bilateral mean FF % was used in the analysis. Baseline visit, follow-up visit, including mapping of ROIs was performed by the same examiner.

We analyzed our MRI data in two parts: First analysis consisted of all participants ($N = 16$). Based on the total mean muscle FF % per participant, we divided the participants in two groups that were analyzed separately: One group consisted of lesser-affected participants ($N = 5$, FF % $< 35\%$) and the other of more severely affected participants ($n = 11$, FF % $> 35\%$). In addition, based on direct observation, the participants with FF % $< 35\%$ displayed the least difficulty with ambulation such as rising from chair and walking, while participants with FF % $> 35\%$ displayed much greater ambulation difficulty.

Statistical Analysis

Statistical analysis was performed using SPSS v22. All values are mean ± SD, unless otherwise stated. The Mann–Whitney test was used to test the null hypothesis of no difference in FF% from baseline to follow-up and trunk strength from baseline to follow-up. The level of significance was set at $p \leq 0.05$.

RESULTS

The 16 participants with genetically verified Becker muscular dystrophy were 37.1 ± 11.9 years old and had a BMI of 27.5 ±

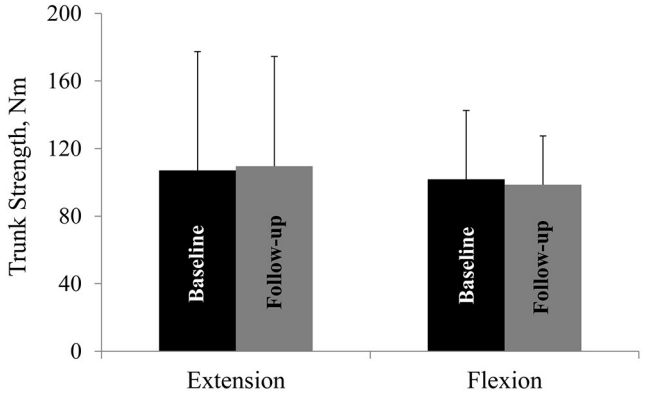


FIGURE 2 | Trunk muscle strength at baseline and at follow-up. Trunk extension and flexion strength at baseline and at follow-up. Nm, Newton-meter. All values are mean ± SD.

6.3 at follow-up. The mean follow-up interval was 398.4 ± 32.1 days (range of 365–478 days).

Trunk Muscle Strength From Baseline to Follow-Up

Mean trunk muscle strength did not differ from baseline to follow-up in the whole cohort (**Figure 2**) or the two subgroups.

Fat Fraction From Baseline to Follow-Up

The first analysis showed no significant change in muscle FF% of the measured muscles of the back, thighs, and lower legs from baseline to follow-up, except for multifidus at spinal level T12 (**Figures 3A–C**), but all muscles showed a nominal increase in fat fraction from baseline to follow-up (3A–C). Abdominal muscles did not show any progression in fat replacement visually inspected at follow-up (**Figure 1**). The subgroup analysis showed a significant increase in FF% in the less affected group at L4/L5 ES ($p = 0.047$), SA ($p = 0.028$), GR ($p = 0.009$), TA ($p = 0.047$), PER ($p = 0.028$), and GM ($p = 0.009$) while the severely affected group increased significantly only at T12 M ($p = 0.028$) and T12 ES ($p = 0.011$).

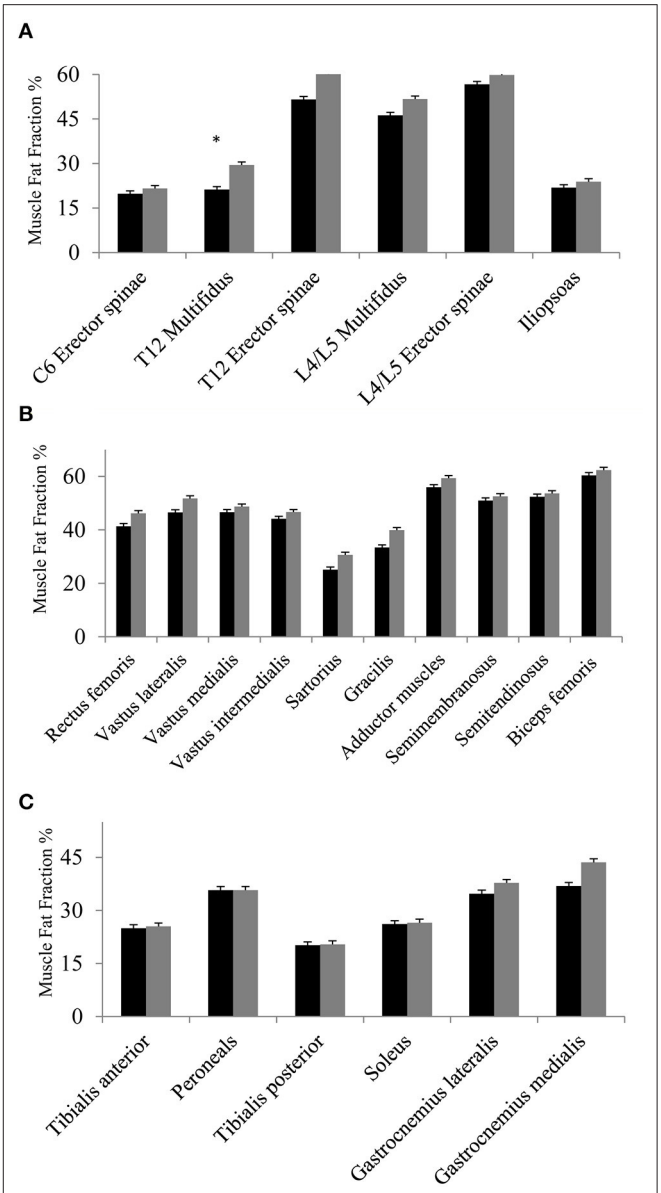


FIGURE 3 | (A–C) Fat fraction % at baseline and at follow-up. Fat fraction % of muscles of the back and iliopsoas **(A)**, muscles of the thighs **(B)**, and muscles of lower legs **(C)** at baseline (black bars) and at follow-up (gray bars). Missing value for spinal level C6 ($n = 1$, due to phase-shift artifact), and missing value for thighs and lower legs ($n = 1$, due to positioning difficulty in scanner). Asterisk * indicates a significant difference ($p = 0.01$). All values are mean ± SD.

DISCUSSION

This is the first study to examine longitudinal changes in disease progression in BMD using qMRI of the paraspinal muscles, lower limb muscles, and trunk muscle strength assessed by dynamometry. We found no significant change in trunk extension and flexion strength from baseline to follow-up, and qMRI on group level only identified a significant increase in FF% in the multifidus at spinal level T12. However, when we separately analyzed the lesser severe and severely affected participants, the FF% increased significantly in the lower back, sartorius, gracilis, and some of the lower leg muscles in the lesser affected individuals, while the severely affected participants only increased in FF% in their paraspinal muscle at spinal level T12. Irrespective of the significant changes in muscle FF%, muscle strength did not differ in the two subgroups.

We quantified trunk extension and flexion strength with a Biodex dynamometer. A previous study has established that strength measure reliability is excellent for trunk extension and moderate for trunk flexion (7). The absence of any change in trunk strength at follow-up, suggests that qMRI, which showed numerical increase in muscle FF%, is a more sensitive marker of change than strength. This discrepancy between qMRI progression and maintained muscle strength is in line

with findings in other muscular dystrophies (8, 9) and is likely explained by the fact that the paraspinal muscles work as a whole unit and that muscle strength of a subset of the paraspinal musculature cannot be measured clinically.

On group level, we did not find any significant difference in FF% from baseline to follow-up, except at spinal level T12 multifidus. In contrast, a significant difference in muscle FF% has been found in 1-year follow-up of other muscular dystrophies (8, 9). One study found a significant difference in all measured muscles of the back, hip flexors, knee extensors and flexors, and lower leg muscles at 1-year follow-up in patients with facioscapulohumeral muscular dystrophy (8), and Willis et al. found a significant increase in FF% in 9 of 14 investigated muscles, including the hamstrings, sartorius, gracilis, vastus lateralis, rectus femoris, and calf muscles in limb–girdle muscular dystrophy type R9 (9). This could suggest that the rate of disease progression in BMD is slower than these other muscular dystrophies and that a 1-year follow-up therefore may be too short to show change. However, our participants had a higher FF% at baseline and therefore there could have been a ceiling effect for change in FF%. In support of this notion, the lesser-affected group of BMD participants significantly increased their muscle FF% in 6 of 22 assessed muscles at follow-up, while the severely affected participants only increased in FF% in one. This suggests that disease progression in BMD may not be much different from FSHD and LGMDR9, when patients with the same disease severity are compared. It also highlights the importance of stratifying patients with BMD according to phenotype for future clinical trials. In addition, the small number of participants in our study may partly explain the lack of significant progression in FF% on a group level.

Another limitation to our study may be the use of 2-point Dixon imaging. Two-point Dixon has shown to be sensitive to phase-shift artifacts in comparison with 3-point Dixon (10); however, we only experienced this with one subject at spinal level C6 and those images were not included in the analysis.

A recent study reported that increases in T2 heterogeneity were observed in fat-replaced muscles in relation to increased disease activity (11). Therefore, it may be of interest to add T2 to MRI protocols in future studies to expand the understanding of disease activity.

In conclusion, we recommend that for future trials patients with BMD should be stratified according to phenotype because they evolve differently and that there is an emphasis on including mild to moderately affected patients, because capturing a change in disease progression in severely affected patients may not be as evident as the lesser affected patients.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Danish National Committee on Health Research Ethics (approval number: H-16030358). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AMS contributed with data collection, data analysis, and drafting of manuscript. KR contributed with study design, data collection, and drafting of manuscript. NW contributed with study design and drafting of manuscript. JV contributed with study design, data analysis, and drafting of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

We thank Grosserer L. F. Foghts Fond, Oda og Hans Fond, Toyota-Fonden, Aase og Ejnar Danielsens Fond, Torkild Steenbecks Legat, Torben and Alice Frimodts Fond and Hede Nielsens Fond for providing grants.

ACKNOWLEDGMENTS

We thank the patients for their participation in this study.

REFERENCES

1. Tasca G, Iannaccone E, Monforte M, Masciullo M, Bianco F, Laschena F, et al. Muscle MRI in Becker muscular dystrophy. *Neuromuscul Disord NMD*. (2012) 22 (Suppl. 2):S100–6. doi: 10.1016/j.nmd.2012.05.015
2. Faridian-Aragh N, Wagner KR, Leung DG, Carrino JA. Magnetic resonance imaging phenotyping of Becker muscular dystrophy. *Muscle Nerve*. (2014) 50:962–7. doi: 10.1002/mus.24246
3. Bonati U, Schmid M, Hafner P, Haas T, Bieri O, Gloor M, et al. Longitudinal 2-point dixon muscle magnetic resonance imaging in becker muscular dystrophy. *Muscle Nerve*. (2015) 51:918–21. doi: 10.1002/mus.24629
4. Maggi L, Moscatelli M, Frangiamore R, Mazzi F, Verri M, De Luca A, et al. Quantitative muscle MRI protocol as possible biomarker in Becker muscular dystrophy. *Clin Neuroradiol*. (2020). doi: 10.1007/s00062-019-00875-0. [Epub ahead of print].
5. Harbo T, Brincks J, Andersen H. Maximal isokinetic and isometric muscle strength of major muscle groups related to age, body mass, height, and sex in 178 healthy subjects. *Eur J Appl Physiol*. (2012) 112:267–75. doi: 10.1007/s00421-011-1975-3
6. Zouita S, Ben Salah FZ, Behm DG, Chaouachi A. Isokinetic trunk strength, validity, reliability, normative data and relation to physical performance and low back pain: a review of the literature. *Int J Sports Phys Ther*. (2020) 15:160–74. doi: 10.26603/ijspt20200160
7. Rudolf K, Sheikh A, Knak K, Witting N, Vissing J. EP.54Assessment of trunk muscle strength in patients with muscular dystrophies using stationary and hand-held dynamometry: a test-retest reliability study. *Neuromuscul Disord*. (2019) 29:S116–7. doi: 10.1016/j.nmd.2019.06.286
8. Andersen G, Dahlqvist JR, Vissing CR, Heje K, Thomsen C, Vissing J. MRI as outcome measure in facioscapulohumeral muscular

dystrophy: 1-year follow-up of 45 patients. *J Neurol.* (2017) 264:438–47. doi: 10.1007/s00415-016-8361-3
9. Willis TA, Hollingsworth KG, Coombs A, Sveen M-L, Andersen S, Stojkovic T, et al. Quantitative muscle MRI as an assessment tool for monitoring disease progression in LGMD2I: a multicentre longitudinal study. *PLoS ONE.* (2013) 8:e70993. doi: 10.1371/journal.pone.0070993
10. Coombs BD, Szumowski J, Coshow W. Two-point Dixon technique for water-fat signal decomposition withB0 inhomogeneity correction. *Magn Reson Med.* (1997) 38:884–9. doi: 10.1002/mrm.1910380606
11. Hooijmans MT, Froeling M, Koeks Z, Verschuuren JJGM, Webb A, Niks EH, et al. Multi-parametric MR in Becker muscular dystrophy patients. *NMR Biomed.* (2020) 33:e4385. doi: 10.1002/nbm.4385

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer GT declared a past co-authorship with one of the authors JV to the handling Editor.

frontiers in Neurology

ORIGINAL RESEARCH
published: 04 February 2021
doi: 10.3389/fneur.2020.621202

Dual Energy Computed Tomography of Internal Carotid Artery: A Modified Dual-Energy Algorithm for Calcified Plaque Removal, Compared With Digital Subtraction Angiography

Hongying Qu[1,2], *Yongan Gao*[1,2], *Meiling Li*[1,2], *Shuo Zhai*[1,2], *Miao Zhang*[1,2] *and Jie Lu*[1,2,3]*

[1] Department of radiology, Xuanwu Hospital, Capital Medical University, Beijing, China, [2] Beijing Key Laboratory of Magnetic Resonance Imaging and Brain Informatics, Beijing, China, [3] Department of Nuclear Medicine, Xuanwu Hospital, Capital Medical University, Beijing, China

Background: Atherosclerotic disease of the internal carotid artery (ICA) is a common reason for ischemic stroke. Computed tomography angiography (CTA) is a common tool for evaluation of internal carotid artery (ICA) stenosis. However, blooming artifacts caused by calcified plaques might lead to overestimation of the stenosis grade. Furthermore, the intracranial ICA is more vulnerable to calcification than other ICA segments. The proposed technique, dual-energy computed tomography (DECT) with a modified three-material decomposition algorithm may facilitate the removal of calcified plaques and thus increase diagnostic accuracy.

Objectives: The objective of the study is to assess the accuracy of the modified three-material decomposition algorithm for grading intracranial ICA stenosis after calcified plaque removal, with digital subtraction angiography (DSA) used as a reference standard.

Materials and Methods: In total, 41 patients underwent DECT angiography and DSA. The three-material decomposition DECT algorithm for calcium removal was applied. We evaluated 64 instances of calcified stenosis using conventional CTA, the previous non-modified calcium removal DECT technique, the modified DECT algorithm, and DSA. The correlation coefficient (r^2) between the results generated by the modified algorithm and DSA was also calculated.

Results: The virtual non-calcium images (VNCa) produced by the previous non-modified calcium removal algorithm were named VNCa 1, and those produced by the modified algorithm were named VNCa 2. The assigned degree of stenosis of VNCa 1 (mean stenosis: 39.33 ± 19.76%) differed significantly from that of conventional CTA images (mean stenosis: 59.03 ± 25.96%; $P = 0.001$), DSA (13.19 ± 17.12%, $P < 0.001$). VNCa 1 also significantly differed from VNCa 2 (mean stenosis: 15.35 ± 18.70%, $P < 0.001$). In addition, there was a significant difference between the degree of stenosis of VNCa 2 and conventional CTA images ($P < 0.001$). No significant differences were observed between VNCa 2 and DSA ($P = 0.076$). The correlation coefficient (r^2) between the stenosis degree of the VNCa 2 and DSA images was 0.991.

OPEN ACCESS

Edited by:
Anna Pichiecchio, Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy

Reviewed by:
Bo Gao, Affiliated Hospital of Guizhou Medical University, China
Stefano Bastianello, University of Pavia, Italy

***Correspondence:**
Jie Lu
smile19761989@163.com

Specialty section:
This article was submitted to Applied Neuroimaging, a section of the journal Frontiers in Neurology

Received: *25 October 2020*
Accepted: *22 December 2020*
Published: *04 February 2021*

Citation:
Qu H, Gao Y, Li M, Zhai S, Zhang M and Lu J (2021) Dual Energy Computed Tomography of Internal Carotid Artery: A Modified Dual-Energy Algorithm for Calcified Plaque Removal, Compared With Digital Subtraction Angiography. Front. Neurol. 11:621202. doi: 10.3389/fneur.2020.621202

Conclusions: The proposed DECT with a modified three-material decomposition algorithm for calcium removal has high sensitivity for the detection of relevant stenoses, and its results were more strongly correlated with DSA than with those of conventional CTA or the previous non-modified algorithm. Further, it overcomes CTA's previous problem of overestimating the degree of stenosis because of blooming artifacts caused by calcified plaques. It is useful to account for calcified plaques while evaluating carotid stenosis.

Keywords: dual-energy computed tomography, internal carotid artery, calcified plaque, calcified stenosis, blooming artifacts, modified algorithm

INTRODUCTION

The incidence of ischemic cerebrovascular disease has increased alongside the increasing numbers of patients with hypertension, hyperglycemia, and hyperlipidemia (1). It is a major cause of death and disability, and it seriously reduces patients' quality of life (2). Internal carotid artery (ICA) atherosclerotic disease is a common cause of ischemic stroke (3). The intracranial ICA has a narrower lumen than the extracranial ICA, and it is more prone to calcification (4). Thus, it might be more vulnerable to blooming artifacts generated by calcified plaque. Such blooming artifacts of calcification in the intracranial ICA might lead to overestimation of vascular stenosis. Recently, for patients with intracranial ICA stenosis, standard drug treatment has been the first-line therapy. However, ischemic cerebrovascular events still occur frequently in some patients. Invalid patients with symptomatic severe intracranial ICA stenosis are suitable for interventional therapy according to the Warfarin and Aspirin for Symptomatic Intracranial Disease (WASID) study. Thus, it is necessary to evaluate the degree of vascular stenosis so that the optimal treatment can be chosen. Further, choosing between different treatments according to the degree of vascular stenosis is important for patients' prognosis (5, 6).

Digital subtraction angiography (DSA) is the gold standard for evaluation of the degree of ICA stenosis, but it is invasive and has periprocedural risks (7). Compared with DSA, some other methods are less invasive and used more prevalently, such as transcranial color-coded duplex sonography (TCCD), magnetic resonance arteriography (MRA), and computed tomography arteriography (CTA). TCCD relies more on the experience and expertise levels of the sonographer. Hence, that method has low repeatability. Another non-invasive method, CE-MRA or TOF-MRA, is limited by contraindications such as claustrophobia and cardiac pacemakers, and it takes a long time to perform. In addition, patients complicated with dysphoria cannot stay still for a long time during scanning, so it is also not suitable for such patients. Besides, TOF-MRA has low spatial resolution and high vulnerability to artifacts caused by blood flow (8). Another widely accepted technique for evaluating ICA stenosis is CTA, which has the advantages of three-dimensional volumetric data analysis, better visualization of distal arteries, and shorter time cost (9). The results of CTA correlate well with those of DSA, and have high diagnostic accuracy (3). However, CTA may be hampered by calcified plaques, which may cause blooming artifacts and lead to overestimation of vascular stenosis (10). Nevertheless, information about the degree of vascular stenosis can be used to select the best treatment for patients (5, 11).

Dual-energy CT (DECT) has been used with increasing frequency in recent years. This modality allows the simultaneous acquisition of low- and high-energy images in a single examination and thus avoids interscan motion. It also reduces the radiation dose and scan time (12). In previous research, the CT value of iodine increased much more than those of bone and calcification when the X-ray tube voltage was decreased (13, 14). Thus, DECT may also have the ability to distinguish between medium-contrast volumes and calcified plaques and then remove the influence of calcified plaques. Some studies have investigated the usefulness of DECT angiography for removing calcified plaques and bone (15). Uotani et al. performed plaque and bone removal to assess the degree of carotid stenosis and concluded that DE hard plaque removal is useful for the evaluation of ICA stenosis with calcification (16). Werncke et al. applied plaque and bone removal to assess the degree of stenosis of peripheral arteries. They observed that DE hard plaque removal is highly effective for heavily calcified plaques (17). However, when Thomas et al. (5) performed plaque and bone removal on 25 patients' images, the results indicated that although DECT had a stronger correlation with DSA than conventional CTA, it frequently overestimated the degree of stenosis. However, distinguishing the degree of stenosis is clinically necessary (5, 18). Thus, a modified technique is needed. Recently, Mannil et al. used a novel modified DECT material-differentiating algorithm to evaluate the degree of stenosis of the extracranial ICA and found that their algorithm removed calcified plaques more accurately (19). However, patients with severe stenosis were not included (19), and research only applied the modified method to detect stenosis of the extracranial ICA.

Therefore, the purpose of this study is to evaluate the accuracy of the novel modified DECT material-differentiating algorithm for detecting the degree of stenosis of the intracranial ICA.

MATERIALS AND METHODS

Patients

A single-center, prospective trial was conducted from January 2018 to March 2019. We collected patients' data from the Department of Neurosurgery, and all patients' medical history was assessed. Physical examinations, laboratory testing, and imaging examinations were also conducted.

The inclusion criteria were patients with intracranial ICA stenosis who underwent both intracranial and extracranial DSA and DE-CTA, the intracranial ICA stenosis was caused by calcified plaque, and there was a <2-week interval between DSA and DE-CTA. The exclusion criteria were unstable clinical conditions. Each patient signed a written consent form, and this research conformed to the principles outlined in the Declaration of Helsinki. We acquired ethical approval from Xuanwu Hospital of Capital Medical University.

CTA Data Acquisition and Postprocessing

The CTA data were acquired with a 128-row multidetector CT scanner (Somatom Force®, Siemens, Munich, Germany). All patients were scanned from the aortic arch to the supraventricular white matter. The scan parameters were tube voltages of 150 and 90 kV, automatic adjustment of tube current according to patient size, pitch 1, slice thickness of 0.75 mm, layer spacing of 0.4 mm, rotation time of 0.5 s, FOV of 19–22 cm, and matrix size of 512 × 512.

The enrolled patients received 65 ml of contrast (Ultravist 370®, BayerSchering Pharma, Berlin, Germany) at a flow rate of 5 ml/s. The precise timing of the injection was determined using a test-bolus technique.

The conventional mixed CTA images were reconstructed with a weighting factor of 0.5, which resulted in a mixture of images from the 150- and 90-kV scans that resembled single-energy CTA. Then, the CTA data were transferred to a workstation (syngofastView®, Siemens Healthineers, Erlangen, Germany) for postprocessing. Virtual non-calcium (VNCa) images were generated according to the previous non-modified calcium removal DECT technique (VNCa 1) and the modified DECT algorithm (VNCa 2) using dedicated prototype software (eXamine, Version 0.9.10; Siemens).

Qualitative Evaluation

Two experienced radiologists evaluated the results, and each was blind to the other's results. The image quality of conventional CTA and VNCa images was assessed on the following scale: 1 = perfect; 2 = not perfect, but could diagnose; 3 = non-diagnostic.

Quantitative Evaluation

The stenosis measurements of the intracranial ICA on DSA, conventional CTA, and VNCa images were evaluated according to the criteria established by the North American Symptomatic Carotid Endarterectomy Trial. The formula: (the diameter of the normal artery beyond the stenosis–the diameter of the narrowest lumen)/the diameter of the normal artery beyond the stenosis × 100%.

Mean intraluminal attenuation values (Hounsfield units, HU) were measured in matched locations on conventional CTA and VNCa 2 images using a standardized circled region of interest. Two different experienced radiologists evaluated the DSA results, and they were blind to the CTA results.

Statistical Analysis

Continuous variables were expressed as means ± standard deviations, and categorical variables were expressed as frequencies or percentages. Differences between stenosis measurements on conventional CTA, VNCa, and DSA were tested using analyses of variance. The differences in mean intraluminal attenuation values between VNCa 2 and conventional CTA were compared by a paired-samples *t*-test. Then, Pearson correlation was used to assess the relationship between DSA and VNCa 2. $P < 0.05$ were considered significant. Intrareader and interreader agreement regarding the degree of stenosis and qualitative imaging parameters was determined by intraclass correlation coefficients and Goodman and Kruskal's gamma, respectively.

TABLE 1 | Descriptive statistics.

			P
Mean age, years	ALL	56 ± 7	
Sex	Male	29 (70.73%)	
	Female	12 (29.27%)	
Side	Left side	33 (51.56%)	
	Right sidetab	31 (48.44%)	
Good diagnostic image quality (%)	Computed tomography angiography (CTA)	62 (96.88%)	
	Virtual non-calcium images produced by the previous non-modified calcium removal algorithm (VNCa 1) and 2	62 (96.88%)	
Mean attenuation, HU	CTA	435.30 ± 83.60	0
	Virtual non-calcium images produced by the modified algorithm (VNCa 2)	399.37 ± 84.24	
Mean stenosis, %	CTA	59.03 ± 25.96	0.001
	VNCa 1	39.33 ± 19.76	
	VNCa 2	15.35 ± 18.70	
	Digital subtraction angiography (DSA)	10.78 ± 2.44	

RESULTS

This prospective study included data from 41 patients (29 men, 12 women) with 64 instances of calcified plaque on the intracranial ICA. Detailed information is found in **Table 1**.

Qualitative Analysis

The image quality of the conventional CTA and VNCa was almost perfect. Only one patient's image with two instances of calcified plaque on the intracranial ICA was not perfect, and that result was influenced by a metal artifact. So in the next analysis, we removed the patient's data.

Quantitative Analysis

The mean levels of carotid artery stenosis in DSA, conventional CTA, and VNCa 1 and 2 images were 13.79 ± 17.12%, 59.03 ± 25.96%, 39.33 ± 19.76%, and 15.35 ± 18.70%, respectively. Significant differences in stenosis level were observed among these four methods ($P = 0.001$; **Table 1**). The VNCa 1 results showed significant differences from those of conventional CTA, DSA images. VNCa 1 also significantly differed from VNCa 2. In addition, the stenosis level in conventional CTA images was significantly different from that in VNCa 2 and DSA images. There was no difference in stenosis levels between VNCa 2 and DSA images ($P = 0.076$; **Table 2**, **Figures 1**, **2**). Representative examples of calcified carotid stenosis in DSA, conventional CTA, and VNCa 1 and 2 images are depicted in **Figures 3**, **4**.

TABLE 2 | Comparison of CTA, VNCa, and DSA stenosis measurements.

Stenosis 1, %	Mean 1 ± SD	Stenosis 2, %	Mean 2 ± SD	P (Stenosis 1 vs. 2)
VNCa 1	39.33 ± 19.76	CTA	59.03 ± 25.96	0.001
		VNCa 2	15.35 ± 18.70	0.000
		DSA	13.19 ± 17.12	0.000
VNCa 2	15.35 ± 18.70	CTA	59.03 ± 25.96	0.000
		DSA	13.19 ± 17.12	0.076
CTA	59.03 ± 25.96	DSA	13.19 ± 17.12	0.000

The mean intraluminal attenuation on conventional CTA images was 435.3 ± 83.6 HU, and the corresponding value at matched locations on VNCa 2 images was 399.37 ± 84.24 HU. The mean intraluminal attenuation on VNCa 2 images was significantly lower than that on conventional CTA images ($P < 0.001$; **Table 1**).

We selected four examples of DSA images for stenosis quantification, and they were evaluated as 30, 90, 75, and 50% stenosis. The same four lesions were quantified as 70, 100, 90, and 70% stenosis, respectively, on conventional CTA images; 50, 100, 90, and 50%, respectively, on VNCa 1 images; and 40, 90, 80, and 50%, respectively, on VNCa 2 images (**Table 3**). Among these lesions, conventional CTA and VNCa 1 tended to overestimate severe stenoses as occlusions or overestimate mild stenoses as moderate–severe stenoses. However, VNCa 2 and DSA had good consistency. The correlation coefficient (r^2) of stenosis grading between the VNCa 2 and DSA images was 0.991.

DISCUSSION

Our results indicate that the modified DECT algorithm is capable of removing calcified plaques and bone. Thus, it achieves improved quantification of intracranial ICA stenosis with DSA as the reference standard.

One of the innovations of DECT is that it overcomes mis-registration, which may occur in the conventional CTA method

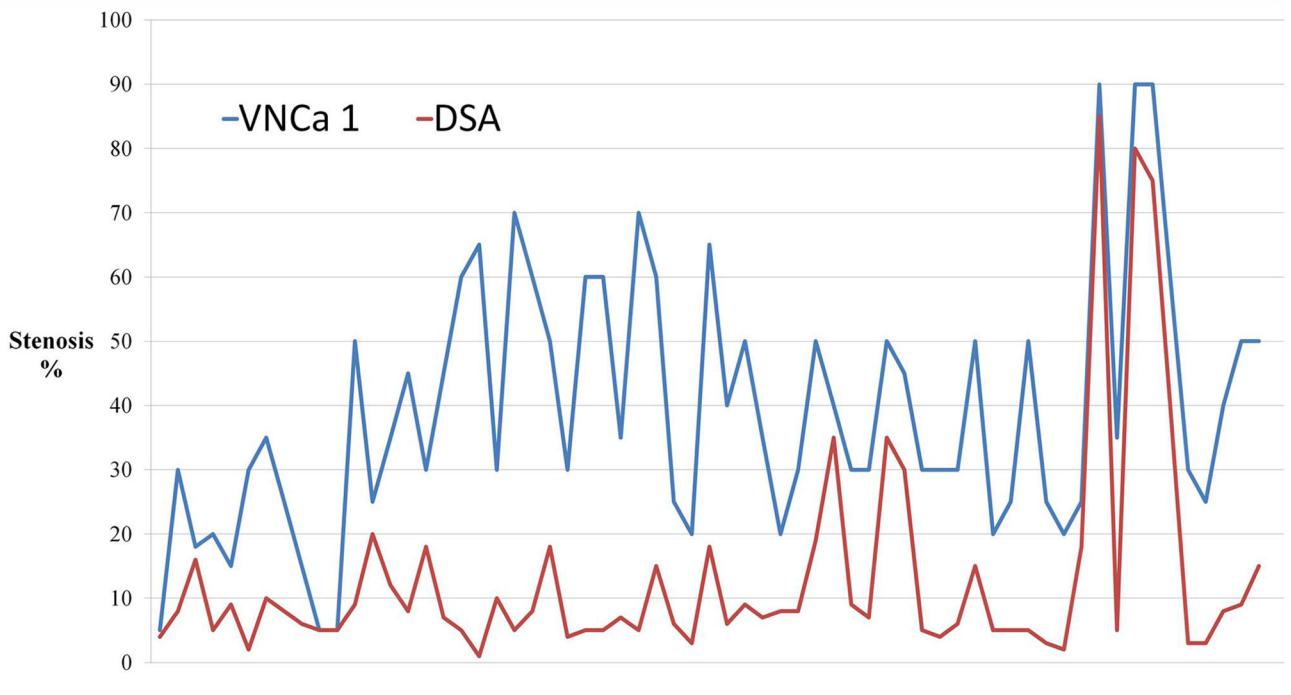


FIGURE 1 | The dot-line graph showed the degree of calcified stenosis in virtual non-calcium images produced by the previous non-modified calcium removal algorithm (VNCa 1) tend to overcome the degree of calcified stenosis compared with digital subtraction angiography (DSA).

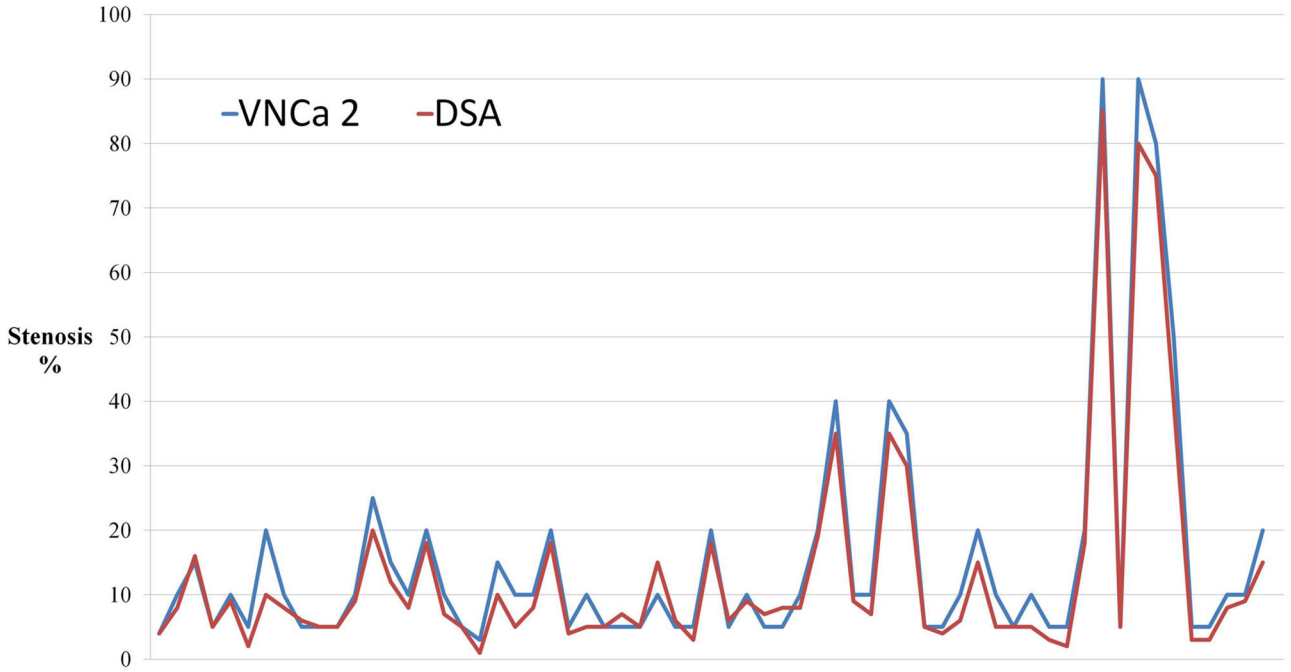


FIGURE 2 | The dot-line graph showed that virtual non-calcium images produced by the modified algorithm (VNCa 2) had good consistency with DSA. Novel three-material decomposition dual-energy computed tomography (DECT) algorithm improves the diagnostic accuracy of computed tomography angiography (CTA).

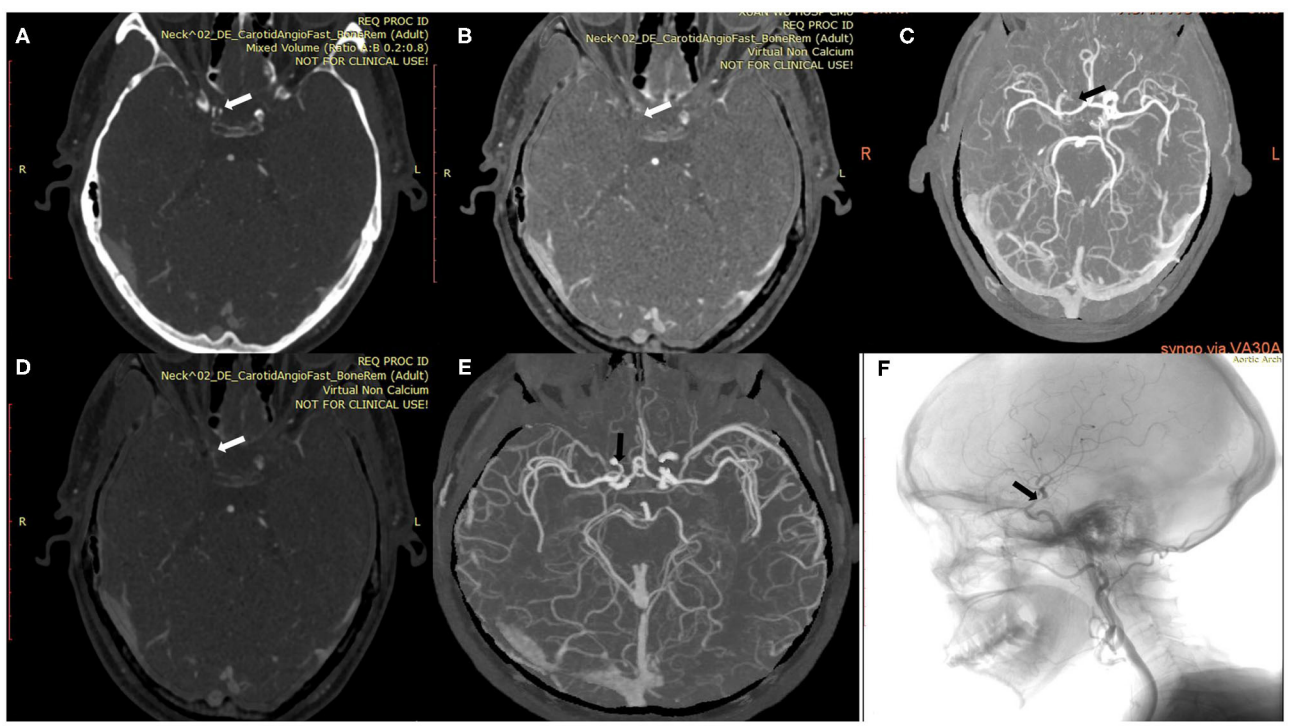


FIGURE 3 | A 54-year-old man with calcified severe stenosis on the right intracranial internal carotid artery (ICA). **(A)** Axial conventional CTA at the location of calcified carotid artery stenosis (arrow). **(B,C)** Axial VNCa 1 and MIP at the location of calcified carotid artery stenosis (arrow). **(D,E)** Axial VNCa 2 and MIP at the location of calcified carotid artery stenosis (arrow). **(F)** DSA in sagittal projection indicated the severe stenosis (arrow). VNCa 2 had good consistency with DSA, and VNCa 1 tend to overcome the severe stenosis to occlusion.

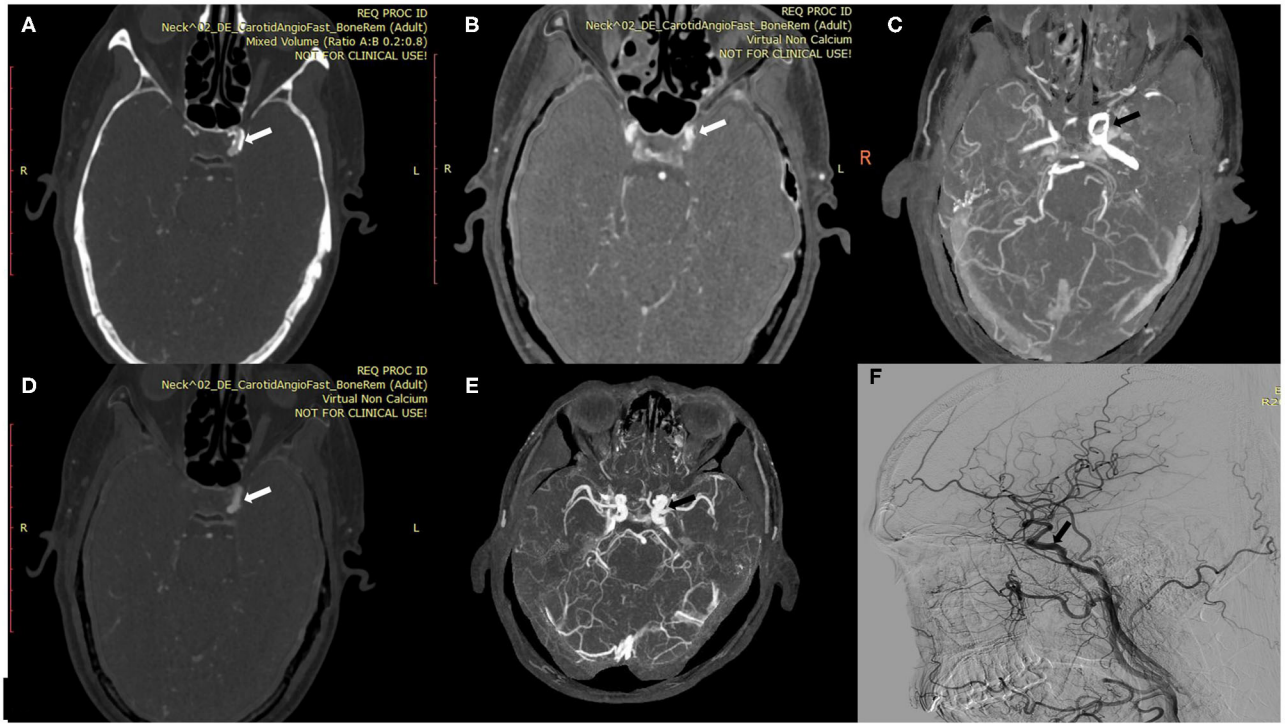


FIGURE 4 | A 60-year-old man with calcified plaque on the left intracranial ICA. **(A)** Axial conventional CTA at the location of calcified plaque (arrow). **(B,C)** Axial VNCa 1 and MIP at the location of calcified carotid plaque (arrow). **(D,E)** Axial VNCa 2 and MIP at the location of calcified plaque (arrow). **(F)** DSA in sagittal projection indicated the no stenosis at the location of calcified plaque (arrow). The VNCa 1 still had residual calcification, and it tends to overcome the degree of calcified stenosis. Then VNCa 2 had good consistency with DSA.

TABLE 3 | Patients with moderate to severe stenosis on DSA.

Patients	1	2	3	4
Age, years	49	61	75	76
Sex	Male	Male	Male	Female
The location of severe stenosis	C6 segment of right ICA	C6 segment of right ICA	C6 segment of left ICA	C4 segment of Right ICA
Degree of stenosis on DSA(%)a	30	90	75	50
Degree of stenosis on CTA(%)a	70	100	90	70
Degree of stenosis on VNCa 1(%)a	50	100	90	50
Degree of stenosis on VNCa 2(%)a	40	90	80	50

as a result of pulsations or neck movements during pre- and post-contrast scanning (20, 21). It has the advantages of low radiation dose and short scan time (22, 23), and it can remove the influence of calcified plaques and increase the accuracy of diagnosis of the degree of vascular stenosis.

Some studies have investigated the usefulness of the early DECT algorithm for removing calcified plaques and bone. For example, Uotani et al. (16) and Werncke et al. (17) applied DE plaque and bone removal to assess the degree of stenosis of carotid or peripheral arteries, finding that the technique was highly effective for heavily calcified plaques. However, some studies have also found that this technique has a disadvantage: it frequently leads to overestimation of the degree of vascular stenosis (5). Standard drug treatment is the first-line therapy for patients with intracranial ICA stenosis. However, some patients with severe intracranial ICA stenosis might fail to respond to

medication, and interventional therapy could be suited for such patients, according to the WASID study. Further, interventional techniques may be unsuitable for patients with occlusion of the intracranial ICA. Hence, distinguishing the degree of stenosis is clinically important to choose suitable treatments for patients (5). Because of the above disadvantage, Mannil et al. applied a modified DECT algorithm to evaluate stenosis in the extracranial ICA and indicated that this new algorithm might improve the technique's accuracy (19). However, that study neglected to investigate intracranial ICA stenosis and did not compare the modified DECT algorithm with previous calcium removal DECT techniques.

In the present study, we applied this modified DECT algorithm to evaluate the degree of stenosis of the intracranial ICA. The intracranial ICA has a smaller lumen than the extracranial ICA. The tortuosity of the intracranial ICA makes it more prone to calcification of the vessel walls (24, 25). The influence of blooming artifacts in the intracranial ICA is more obvious than that in the extracranial ICA. The intracranial ICA is located at the base of the skull. Sometimes, the skull base bone can also interfere with judgment of vascular stenosis (26). Our results indicate that the modified DECT algorithm is more accurate than conventional CTA and previous non-modified calcium removal DECT technique, and its results are strongly correlated with those of DSA. This modified DECT algorithm has the potential to effectively increase the clinicians' ability to test the degree of vascular stenosis. Besides, the previous non-modified calcium removal DECT technique had higher diagnostic accuracy than conventional CTA but also tended to overestimate the degree of vascular stenosis.

In addition, we enrolled some patients with moderate–severe stenosis of the intracranial ICA in this study, providing a necessary complement to previous studies. Distinguishing between severe stenosis and occlusion or between mild–moderate stenosis and severe stenosis is clinically important, as the treatments for those conditions are different. For example, occlusions are not appropriate for stenting. If treatable stenoses are categorized as occlusions, treatment could be denied to patients (5, 18). Further, interventional therapy is not recommended for mild–moderate intracranial ICA stenosis. Thus, the assessment of the degree of vascular stenosis could affect patients' prognosis. In this study, the modified DECT algorithm and DSA produced consistent results, but conventional CTA and the previous non-modified calcium removal DECT technique tended to overestimate vascular stenosis. Thus, the modified DECT algorithm could improve the clinicians' diagnostic ability.

In this study, the modified DECT algorithm successfully removed calcified plaques from the intracranial ICA in almost all cases (96.88%). Only two calcified plaques were removed insufficiently (i.e., large residual calcifications were present). Serious metal artifacts might lead to undesirable results. Also, the level of intraluminal attenuation on VNCa 2 was significantly lower than that on conventional CTA images, indicating that the modified DECT calcium removal algorithm could partially remove iodine as well, which implies that the modified algorithm still has limitations in its current form.

Our study has the limitation of a small sample size. Larger trials need to be conducted to test our results.

The novel three-material decomposition DECT algorithm removed the calcified plaques on intracranial ICA in CTA images effectively and improved image quality. Its results were strongly correlated with those of DSA and overcame CTA's previous problem of overestimating degree of vascular stenosis. Thus, it has the potential to improve the diagnostic accuracy of CTA.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Xuanwu Hospital of Capital Medical University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HQ participated in the experimental design, data collection and post-processing, and paper writing. ML and SZ participated in the data collection and post-processing. JL not only provided the experimental equipment and post processing workstation but also revised the first draft of the paper. MZ participated in the data collection and giving a hand in the statistical analysis. YG participated in the experimental design, data post-processing, and revising of the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Beijing Municipal Administration of Hospitals' Ascent Plan (Code: DFL20180802).

REFERENCES

1. Rennenberg RJ, Kessels AG, Schurgers LJ, van Engelshoven JM, de Leeuw PW, Kroon AA. Vascular calcifications as a marker of increased cardiovascular risk: a meta-analysis. *Vasc Health Risk Manag.* (2009) 14:185–97. doi: 10.2147/VHRM.S4822

2. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. AHA statistical update heart disease and stroke statistics-2016 update. *Circulation.* (2016) 40:e38–360. doi: 10.1161/CIR.0000000000000350

3. Ertl-Wagner B, Brüning R, Hoffmann RT, Meimarakis G, Reiser MF. Diagnostic evaluation of carotid artery stenoses with multislice CT

angiography: review of the literature and results of a pilot study. *Radiologe.* (2004) 10:960–6. doi: 10.1007/s00117-004-1108-7
4. Babiarz LS, Yousem DM, Wasserman BA, Wu C, Bilker W, Beauchamp NJJ. Cavernous carotid artery calcification and white matter ischemia. *Am J Neuroradiol.* (2003) 24:872.
5. Thomas C, Korn A, Ketelsen D, Danz S, Tsifikas I, Claussen CD, et al. Automatic lumen segmentation in calcified plaques: dual-energy CT vs. standard reconstructions in comparison with digital subtraction angiography. *Am J Roentgenol.* (2010) 194:1590–5. doi: 10.2214/AJR.09.3550
6. Jang E, Chung J, Seo K, Suh SH, Kim YB, Lee K. A protocol-based decision for choosing a proper surgical treatment option for carotid artery stenosis. *J Cerebrovasc Endovasc Neurosurg.* (2015) 17:101–4. doi: 10.7461/jcen.2015.17.2.101
7. Josephson SA. Evaluation of carotid stenosis using CT angiography in the initial evaluation of stroke and TIA. *Neurology.* (2004) 63:457–60. doi: 10.1212/01.WNL.0000135154.53953.2C
8. Francesca MC, Joanna MW, Fmed S. Carotid artery stenosis: accuracy of noninvasive tests -individual patient data meta-analysis. *Neurology.* (2009) 251:493–502. doi: 10.1148/radiol.2512080284
9. Albrecht T, Foert E, Holtkamp R, Kirchin MA, Ribbe C, Wacker FK, et al. 16-MDCT angiography of aortoiliac and lower extremity arteries: comparison with digital subtraction angiography. *Am J Roentgenol.* (2007) 189:702–6. doi: 10.2214/AJR.07.2333
10. Khan M, Qureshi AI. Factors associated with increased rates of post-procedural stroke or death following carotid artery stent placement: a systematic review. *J Vasc Interv Neurol.* (2014) 7:11–8.
11. Johnson TR, Krauss B, Sedlmair M, Grasruck M, Bruder H, Morhard D, et al. Material differentiation by dual energy CT: initial experience. *Euro Radiol.* (2007) 17:1510–7. doi: 10.1007/s00330-006-0517-6
12. Gert JP, Robbert W, Van HM. Accuracy of iodine quantification using dual energy CT in latest generation dual source and dual layer CT. *Eur Radiol.* (2017) 27:3902–14. doi: 10.1007/s00330-017-4752-9
13. Nishizawa Y, Higuchi C, Nakaoka T, Omori H, Ogawa T, Sakura H, et al. Compositional analysis of coronary artery calcification in dialysis patients *in vivo* by dual-energy computed tomography angiography. *Ther Apher Dial.* (2018) 22:365–70. doi: 10.1111/1744-9987.12662
14. Kaemmerer N, Brand M, Hammon M, May M, Wuest W, Krauss B, et al. Dual-energy computed tomography angiography of the head and neck with single-source computed tomography. *Invest Radiol.* (2016) 51:618–23. doi: 10.1097/RLI.0000000000000290
15. Uotani K, Watanabe Y, Higashi M, Nakazawa T, Kono AK, Hori Y, et al. Dual-energy CT head bone and hard plaque removal for quantification of calcified carotid stenosis: utility and comparison with digital subtraction angiography. *Eur Radiol.* (2009) 19:2060–5. doi: 10.1007/s00330-009-1358-x
16. Werncke T, Albrecht T, Wolf K, Meyer BC. Dual energy CT of the peripheral arteries: a phantom study to assess the effect of automatic plaque removal on stenosis grading. *Rofo.* (2010) 182:682–9. doi: 10.1055/s-0029-1245268
17. Fanous AA, Natarajan SK, Jowdy PK, Dumont TM, Mokin M, Yu J, et al. High-risk factors in symptomatic patients undergoing carotid artery stenting with distal protection: buffalo risk assessment scale. *Neurosurgery.* (2015) 4:531–43. doi: 10.1227/NEU.0000000000000871
18. Mannil M, Ramachandran J, Vittoria de Martini I, Wegener S, Schmidt B, Flohr T, et al. Modified dual-energy algorithm for calcified plaque removal. *Invest Radiol.* (2017) 52:680–5. doi: 10.1097/RLI.0000000000000391
19. Den H, Henk WVF. CT Angiography of the circle of willis and intracranial internal carotid arteries: maximum intensity projection with matched mask bone elimination—feasibility study. *Radiology.* (2001) 213:893–8. doi: 10.1148/radiology.218.3.r01mr30893
20. Lell MM, Ditt H, Panknin C, Sayre JW, Ruehm SG, Klotz E, et al. Bone-subtraction CT angiography: evaluation of two different fully automated image-registration procedures for interscan motion compensation. *Am J Neuroradiol.* (2007) 28:1362–8. doi: 10.3174/ajnr.A0558
21. Daniel S, Jit P. Radiation dose efficiency of dual-energy CT benchmarked against single source, kilovoltage-optimized scans. *Br Instit Radiol.* (2016) 89:2015–22. doi: 10.1259/bjr.20150486
22. Agostini A, Mari A, Lanza C, Schicchi N, Borgheresi A, Maggi S, et al. Trends in radiation dose and image quality for pediatric patients with a multidetector CT and a third-generation dual-energy CT. *Radiol Med.* (2019) 124:745–52. doi: 10.1007/s11547-019-01037-5
23. Damaskos S, Aartman IHA, Tsiklakis K, van der Stelt P, Berkhout WER. Association between extra- and intracranial calcifications of the internal carotid artery: a CBCT imaging study. *Dentomaxillofac Radiol.* (2015) 44:2014–8. doi: 10.1259/dmfr.20140432
24. Zhao DL, Wan Y, Wang GK, Wang HB, Liang HW, Zhou HT, et al. Evaluation of image quality in carotid and cerebrovascular disease: a comparative study between subtraction and routine computed tomography angiography. *Echocardiography.* (2016) 33:1735–40. doi: 10.1111/echo.13326
25. May MS, Wiesmueller M, Heiss R, Brand M, Bruegel J, Uder M, et al. Comparison of dual- and single-source dual-energy CT in head and neck imaging. *Eur Radiol.* (2019) 29:4207–14. doi: 10.1007/s00330-018-5762-y
26. Korn A, Bender B, Brodoefel H, Hauser TK, Danz S, Ernemann U, et al. Grading of carotid artery stenosis in the presence of extensive calcifications: dual-energy CT angiography in comparison with contrast-enhanced MR angiography. *Clin Neuroradiol.* (2015) 25:33–40. doi: 10.1007/s00062-013-0276-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 05 February 2021
doi: 10.3389/fneur.2021.625565

Wrist Circumference-Dependent Upper Limit of Normal for the Cross-Sectional Area Is Superior Over a Fixed Cut-Off Value in Confirming the Clinical Diagnosis of Carpal Tunnel Syndrome

Tom B. G. Olde Dubbelink[1], Floriaan G. C. M. De Kleermaeker[2], Roy Beekman[3], Juerd Wijntjes[4], Ronald H. M. A. Bartels[5], Jan Meulstee[1] and Wim I. M. Verhagen[1]*

[1] Department of Neurology, Canisius-Wilhelmina Hospital, Nijmegen, Netherlands, [2] Department of Neurology, VieCuri Medical Centre, Venlo, Netherlands, [3] Department of Neurology, Zuyderland Medical Centre, Heerlen, Netherlands, [4] Department of Neurology, Radboud University Medical Centre, Nijmegen, Netherlands, [5] Department of Neurosurgery, Radboud University Medical Centre, Nijmegen, Netherlands

OPEN ACCESS

Edited by:
Francesco Santini, University of Basel, Switzerland

Reviewed by:
Paolo Felisaz, Azienda Ospedaliera Fatebenefratelli e Oftalmico, Italy
Fabrizio Calliada, University of Pavia, Italy

***Correspondence:**
Tom B. G. Olde Dubbelink
t.oldedubbelink@cwz.nl

Specialty section:
This article was submitted to Applied Neuroimaging, a section of the journal Frontiers in Neurology

Received: *03 November 2020*
Accepted: *15 January 2021*
Published: *05 February 2021*

Citation:
Olde Dubbelink TBG, De Kleermaeker FGCM, Beekman R, Wijntjes J, Bartels RHMA, Meulstee J and Verhagen WIM (2021) Wrist Circumference-Dependent Upper Limit of Normal for the Cross-Sectional Area Is Superior Over a Fixed Cut-Off Value in Confirming the Clinical Diagnosis of Carpal Tunnel Syndrome. Front. Neurol. 12:625565. doi: 10.3389/fneur.2021.625565

Introduction: In confirming the clinical diagnosis of carpal tunnel syndrome (CTS), ultrasonography (US) is the recommended first diagnostic test in The Netherlands. One of the most important parameters for an abnormal US result is an increase of the CSA of the median nerve at the carpal tunnel inlet. An earlier study showed that a wrist-circumference dependent cut-off for the upper limit of normal of this CSA might be superior to a fixed cut-off of 11 mm^2. In this study we compared three ultrasonography (US) parameters in three large Dutch hospitals.

Methods: Patients with a clinical suspicion of CTS and with reasonable exclusion of other causes of their symptoms were prospectively included. A total number of 175 patients were analysed. The primary goal was to compare the number of wrists with an abnormal US result while using a fixed cut-off of 11 mm^2 (FC), a wrist circumference-dependent cut-off ($y = 0.88 * x-4$, where y = ULN and x = wrist circumference in centimetres; abbreviated as WDC), and an intraneural flow related cut-off (IFC).

Results: The WDC considered more US examinations to be abnormal (55.4%) than the FC (50.3%) did, as well as the IFC (46.9%), with a statistically significant difference of $p = 0.035$ and $p = 0.001$, respectively. The WDC detected 12 abnormal median nerves while the FC did not, and 18 while the IFC did not. The wrist circumference of the patients of these subgroups turned out to be significantly smaller ($p < 0.001$) when compared with the rest of the group.

Conclusion: According to these study results, the wrist-circumference dependent cut-off value for the CSA of the median nerve at the wrist appears to have a higher sensitivity than either a fixed cut-off value of 11 mm^2 or cut-off values based on intraneural flow, and may add most value in patients with a smaller wrist circumference.

Keywords: carpal tunnel syndrome, ultrasonography, cross-sectional area (CSA), median nerve, cut-off

INTRODUCTION

Carpal tunnel syndrome (CTS) is the most common peripheral mononeuropathy with a prevalence ranging from 1 to 6% in the general population (1, 2). CTS is caused by compression of the median nerve as it travels through the carpal tunnel, and can be diagnosed clinically. The symptoms classically include pain and paraesthesia in the territory of the median nerve, increasing during the night, and provocation by flexing or extending the wrist (3). Case history evaluation is the most important part of consultation for the clinical diagnosis of CTS (4). Nevertheless, in The Netherlands, most surgeons require a confirmation by an electrodiagnostic or ultrasound test (5).

As of 2017, ultrasonography (US) is the recommended first diagnostic test in The Netherlands because it is easily accessible and painless. Moreover, US and nerve conduction studies (NCS) have a similar sensitivity and specificity, according to the Dutch CTS guideline (6). In literature, several ultrasonography parameters for confirming the diagnosis CTS are suggested, the most important being (7, 8):

- An increase in the cross-sectional area (CSA) of the median nerve at wrist level;
- Flattening ratio of the median nerve at the hamate level;
- Swelling ratio; increase in the CSA at the wrist level compared to the CSA at distal radius level (9);
- Palmar bowing of the flexor retinaculum;
- And hypervascularisation (10).

In a previous study we found that a wrist circumference-dependent cut-off value of the CSA could lead to increased diagnostic accuracy (11). In this study we used the increase of the CSA of the median nerve at wrist level as parameter and evaluated which of three ultrasonographic cut-off values can confirm the clinical diagnosis of CTS the most accurately. These three parameters include a cut-off value of the maximum CSA based on (1) a fixed cut-off value, (2) a wrist circumference-dependent cut-off value, and (3) an intraneural flow-dependent cut-off value (the presence or absence of increased nerve vascularisation).

In earlier studies (12) we used more rigid clinical criteria for diagnosing CTS. Patients were included if they experienced paraesthesias and/or pain in the median nerve-innervated territory, and two or more of the following clinical signs: (1) nocturnal paraesthesias, (2) aggravation of paraesthesias by driving, holding a book or telephone, and (3) a positive Flick sign. However, in our experience, confirmation of presumed CTS is often required by the clinician for patients who do not fulfil all these criteria. In this study we analysed data from three large Dutch teaching hospitals with less strict, but in daily practice more commonly used, inclusion criteria for CTS, as mentioned in paragraph Study Population and Sonography Assessment. Compared to earlier studies, waking up at night due to the symptoms was not mandatory in this study and we did not specify the different types of aggravating activities.

The primary goal of this study was to investigate in how many patients clinical CTS conformation could be achieved by using these three cut-off values and compare performance of these cut-off values. The clinical diagnosis, as defined in the inclusion criteria in the next section, was used as the gold standard.

Abbreviations: CTS, carpal tunnel syndrome; CSA, cross-sectional area; FC, fixed cut-off value; IFC, intraneural flow related cut-off value; IQ, interquartile range; NCS, nerve conduction studies; ULN, upper limit of normal; US, ultrasonography; WDC, wrist circumference-dependent cut-off value.

MATERIALS AND METHODS

Methods

Patients with a clinical suspicion of CTS were referred by their general practitioners between 2018 and 2020. Neurologists in one Dutch university hospital (Radboud university medical centre, Nijmegen) and two Dutch teaching hospitals [Zuyderland hospital, Heerlen and Canisius-Wilhelmina hospital (CWZ), Nijmegen] took the history and included patients if they met the criteria as mentioned in the next paragraph. Ultrasonography was performed and the cross-sectional area (CSA) for the median nerve was measured in all hospitals. The circumference at the distal wrist crease [affected side(s)], height, weight, age, gender and the duration of symptoms were documented. Ultrasonography studies were performed on the same day for each patient and by experienced electrodiagnostic technicians.

The study was approved by the local ethics committee.

Study Population and Sonography Assessment

Patients were included if they met all of the following inclusion criteria:

- Over 18 years old;
- Paraesthesia (possibly accompanied by hypaesthesia and/or pain) in the territory of the median nerve;
- Aggravation of complaints by certain activities or wrist movements;
- Reasonable exclusion of other causes based on history-taking and examination.

In patients with bilateral complaints only one side was randomly included. We excluded patients with clinical signs of polyneuropathy, previous surgery or trauma to the wrist and bifid median nerves. Also, pregnant patients and patients with a history of rheumatoid arthritis, diabetes mellitus, hereditary neuropathy with liability to pressure palsies, thyroid disease or alcoholism were excluded.

US studies were performed by neurophysiology technicians with at least 5 years of nerve ultrasound experience. The studies were performed with a Hitachi Aloka Arietta 850 ultrasound system in the Canisius-Wilhelmina hospital (5–17 MHz linear array transducer) and in the Zuyderland hospital (5–18 MHz linear array transducer). In the Radboud university medical centre (Radboudumc) a Fujifilm Sonosite Xporte (5–16 MHz linear array transducer) was used. The main settings of the US machine were: frequency 17–18 MHz, acoustic power 100%, deepness 1.5 cm and focus position 2 cm. Patients were examined in a sitting position with their forearm in supination resting on an examination couch. The median nerve ultrasonography was performed in longitudinal and transverse planes. The inner

margin of the hyperechoic rim was outlined by the technicians, as was learned in specialised ultrasound training. The CSA was calculated by the software of the ultrasound system, rounding all measurements to the nearest 0.01 cm^2. Colour Doppler sonography with no extra manual compression was used to depict potential intraneural blood vessels. The power Doppler box was placed over the nerve with the focus point adjusted to the nerve depth. The colour gain was set to the maximum level for higher sensitivity to flow signals. Intraneural flow was defined as pulsatile focal colour flow signals. Ultrasonographic protocol in the Radboudumc included measurements of both distal and proximal carpal tunnel CSA; we included the largest measured CSA only.

A fixed cut-off value of >11 mm^2 for the CSA of the median nerve at the wrist level was compared with wrist circumference-dependent cut-off values and cut-off values based on increased nerve vascularisation in the median nerve. In the Dutch CTS guideline (6) a fixed cut-off of 11 mm^2 for the median nerve CSA at wrist level is mentioned based on a study from Visser et al. (13). The wrist circumference-dependent cut-off value was calculated by an equation, $y = 0.88 * x-4.0$, where y is the upper limit of normal of the CSA and x = wrist circumference in centimetres, as described in a previous study (11). If intraneural flow was present a cut-off value of > 12.4 mm^2 was used, and if absent, the cut-off value was >11.2 mm^2. These cut-off values were based on the CSA upper limit of normal 95th percentiles in a healthy population (14).

For readability purposes we abbreviated the cut-off values in the rest of this paper as FC for the fixed cut-off value, WDC for the wrist circumference dependent cut-off value and IFC for the intraneural flow related cut-off value.

Statistics

Statistical analyses were performed using SPSS Statistics Version 26.0. The type of distribution of the data was checked by performing visual analysis of the histograms, the Kolmogorov–Smirnov test and Q-Q plots. Group comparisons for patient

TABLE 2 | Combined results of ultrasonography of the three hospitals for the fixed cut-off, the wrist circumference-dependent cut-off (WDC) and the cut-off based on intraneural flow (IFC).

Sonography		Total number of wrists (*n* = 175)	Zuyderland (*n* = 71)	Radboudumc (*n* = 41)	CWZ (*n* = 63)
Abnormal	FC	88 (50.3%)	21 (29.6%)	23 (56.1%)	44 (69.8%)
	WDC	97 (55.4%)	29 (40.8%)	25 (61.0%)	43 (68.3%)
	IFC	82 (46.9%)	20 (28.2%)	23 (56.1%)	39 (61.9%)

TABLE 1 | Patient characteristics.

	Zuyderland	Radboudumc	Canisius-Wilhelmina hospital (CWZ)	Total	*p*
Participants (*n*)	71	41	63	175	
Men/women	22 (31.0%)/49 (69.0%)	9 (22.0%)/32 (78.0%)	22 (34.9%)/41 (65.1%)	53 (30.3%)/122 (69.7%)	0.367[a]
Mean age (y, SD)	53.8 (15.3)	57.0 (12.8)	57.3 (16.3)	55.8 (15.2)	0.347[b]
Left/right	35/36	16/25	32/31	83/92	0.461[a]
Median height (cm, IQ range)	168.0 (12.0)	168.0 (9.0)	169.0 (9.0)	168.6 (10.0)	0.962[c]
Median weight (kg, IQ range)	77.5 (22.0)	78.0 (21.0)	78.0 (20.0)	78.0 (21.0)	0.559[c]
Median BMI (kg/m^2, IQ range)	27.1 (5.7)	27.1 (5.5)	27.7 (6.2)	27.4 (6.0)	0.210[c]
Median duration symptoms (months, IQ range)	6.0 (20.5)	6.0 (21.8)	12.0 (21.0)	6.0 (22.0)	0.379[c]
Median wrist circumference (cm, IQ range)	17.0 (2.0)	17.0 (2.0)	16.9 (1.7)	17.0 (1.8)	0.950[c]
Minimum-maximum range wrist circumference in cm	14.0-20.0	14.0-19.5	14.4-20.0	14.0-20.0	
Median left wrist circumference (cm, IQ range)	17.0 (1.2)	17.8 (1.4)	16.8 (1.4)		
Mean right wrist circumference (cm, SD)	17.0 (2.7)	17.0 (1.8)	17.2 (1.8)		
Median CSA wrist (mm^2, IQ range)	10.3 (2.7)	12.3 (4.8)	12.2 (5.4)	11.1 (4.3)	<0.001[c†]
Median CSA wrist left (mm^2, IQ range)	10.3 (2.3)	12.8 (6.5)	12.4 (4.3)		
Median CSA wrist right (mm^2, IQ range)	9.8 (2.6)	12.0 (4.2)	12.0 (6.0)		
Intraneural flow	6 (8.5%)	0	13 (20.6%)	19 (10.9%)	0.001[d]
Intraneural flow left	1 (1.4%)	0	7 (11.1%)	8 (4.6%)	
Intraneural flow right	5 (7.0%)	0	6 (9.5%)	11 (6.3%)	
Intraneural flow absent	65 (91.5%)	41 (100%)	50 (79.4%%)	156 (89.1%)	

[a] *Pearson Chi-Square test.*
[b] *One-way between groups ANOVA analysis.*
[c] *Kruskal-Wallis test.*
[d] *Likelihood ratio Chi-Square test.*
[†] *In the Zuyderland hospital compared to CWZ, and the Zuyderland hospital compared to the Radboudumc, a statistically significant different CSA at the wrist, $p < 0.001$ and $p = 0.042$, respectively, was found. The CSA at the wrist in the Radboudumc compared to CWZ was not statistically significantly different ($p = 0.318$) (Mann–Whitney U-test).*

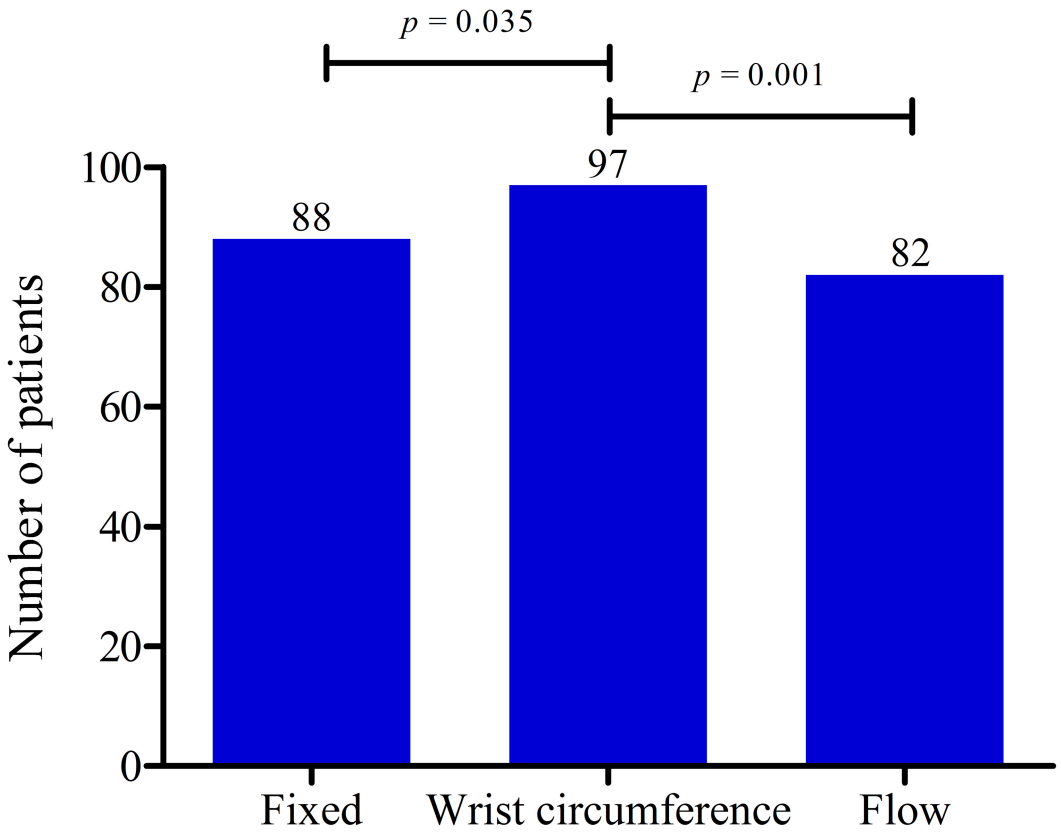


FIGURE 1 | Data of three hospitals combined showing the number of patients with abnormal ultrasonography (US) results. Comparison of the fixed cut-off value (FC), the wrist circumference-dependent cut-off value (WDC) and the flow-dependent cut-off value (IFC). US results were more often abnormal both with the FC ($p = 0.031$) and the WDC ($p = 0.001$) compared with the IFC. *P*-values calculated with McNemar's test.

characteristics were performed by Chi–Square test (nominal unpaired data), one-way ANOVA analysis for numerical, normally distributed, unpaired data and for non-parametric data the Mann–Whitney *U* (2 groups) or Kruskal–Wallis (>2 groups) test. The categorical data of the three hospitals combined was analysed using McNemar's test for paired data. $p < 0.05$ was considered to be statistically significant.

RESULTS

Table 1 shows the patient characteristics. A total number of 175 patients were included. There were no statistically significant differences between the hospitals in gender, age, side of included wrist, height, weight, BMI, or the duration of symptoms. Particularly, wrist circumference did not differ between these patient groups. The CSA of the median nerve at wrist was significantly smaller in the Zuyderland hospital compared with CWZ ($p < 0.001$) and compared with the Radboudumc ($p = 0.042$). This value did not differ between CWZ and the Radboudumc ($p = 0.318$).

Patient Data

In order to assess the primary goals of this study, data of the three hospitals was combined. We compared the total number of abnormal ultrasonography results using the three different parameters as discussed. In **Table 2** the results are presented.

As shown in **Figure 1**, ultrasonography was significantly more often abnormal while using the WDC compared with the FC ($p = 0.035$) or the IFC ($p = 0.001$).

The following paragraphs give insight in the hospital-specific data.

TABLE 3 | Results of ultrasonography in the Zuyderland hospital for the fixed cut-off (FC), the wrist circumference-dependent cut-off (WDC) and the cut-off based on intraneural flow (IFC).

	FC	WDC	IFC
US abnormal	21 (29.6%)	29 (40.8%)	20 (28.2%)
US abnormal left	12	16	12
US abnormal right	9	13	8

Zuyderland Hospital

We collected data from 75 patients. Ten bifid median nerves in nine patients led to four exclusions. We eventually included 71 patients. **Table 3** shows the data of the performed tests of the Zuyderland hospital. Interestingly, ultrasonography was

abnormal in only 28.2% of the examined wrists when using the IFC, while 40.8% were considered abnormal when the WDC was used.

Radboud University Medical Centre

Data from 48 patients was collected. We excluded one patient because of missing data points, two patients with bifid median nerves, two posttraumatic CTS patients and two patients because of prior wrist surgery. Forty-one patients were included. **Table 4** shows the hospital-specific results. 56.1–61.0% of the 41 examined wrists showed abnormal US results.

Canisius-Wilhelmina Hospital (CWZ)

Data of 72 patients was obtained. Seven patients were excluded because of bifid median nerves, Two patients had a medical history of wrist arthrosis. Sixty-three patients were included. In **Table 5** the results of CWZ are presented. Ultrasound test results were abnormal in 61.9–69.8%.

TABLE 4 | Results of ultrasonography in the Radboudumc for the fixed cut-off (FC), the wrist circumference-dependent cut-off (WDC) and the cut-off based on intraneural flow (IFC).

	FC	WDC	IFC
US abnormal	23 (56.1%)	25 (61.0%)	23 (56.1%)
US abnormal left	9	9	9
US abnormal right	14	16	14

TABLE 5 | Results of ultrasonography in the Canisius-Wilhelmina hospital for the fixed cut-off (FC), the wrist circumference-dependent cut-off (WDC) and the cut-off based on intraneural flow (IFC).

	FC	WDC	IFC
US abnormal	44 (69.8%)	43 (68.3%)	39 (61.9%)
US abnormal left	24	22	22
US abnormal right	20	21	17

Comparison of the Ultrasonography Parameters

The WDC considered 12 median nerves to be abnormal while the FC did not, and 18 while the IFC did not. The wrist circumference of these 12 and 18 wrists turned out to be significantly smaller ($p < 0.001$) when compared with the rest of the wrist circumferences as shown in **Table 6**. The mean CSA between these groups was not significantly different. Most of these patients (9/12 in the normal fixed group and 10/18 in the normal flow group) were examined at the Zuyderland hospital.

DISCUSSION

In this cross-sectional study a wrist circumference-dependent cut-off value for the upper limit of normal of the CSA at the wrist led to more abnormal US results than either a fixed upper limit of normal of 11 mm^2 ($p = 0.035$) or a cut-off value based on intraneural flow ($p = 0.001$) did.

As shown in an earlier study, a WDC of the CSA may augment diagnostic accuracy of ultrasonography in CTS patients (11). We found that this parameter considered more US to be abnormal than a FC or an IFC did. The mean CSA of the median nerves of the evaluated patients was relatively low in this study. In the Zuyderland hospital the mean CSA was significantly lower than in the Canisius-Wilhelmina hospital ($p < 0.001$) and the Radboudumc ($p = 0.042$). This lower mean CSA in the Zuyderland hospital leads to a very low number of abnormal US in the Zuyderland hospital, ranging from 28.2 to 40.8%. Looking at the small subgroup of patients with abnormal US results while using the WCD and normal US results while using

TABLE 6 | Patient characteristics of 12 and 18 patients with an abnormal ultrasonography (US) result while using the wrist circumference dependent cut-off value but normal US results using the fixed and the flow-dependent cut-off value, respectively.

	Normal fixed	Others	*p*	Normal flow	Others	*p*
Wrists (*n*)	12	163		18	157	
Men/Women	2/10	51/112	0.287[a]	3/15	50/107	0.184[a]
Mean age (years, SD)	54.5 (12.0)	55.9 (15.4)	0.760[b]	52.8 (12.9)	56.1 (15.4)	0.374[b]
Left/right	5/7	78/85	0.679[a]	7/11	76/81	0.444[a]
Median height (cm, IQ range)	163.0 (11.8)	168.0 (10.3)	0.268[c]	164.0 (11.0)	169.0 (10.8)	0.061[c]
Median weight (kg, IQ range)	69.5 (15.3)	78.5 (20.0)	0.013[c]	72.0 (16.3)	78.5 (20.3)	0.071[c]
Median BMI (kg/m^2, IQ range)	25.4 (4.1)	27.6 (6.0)	0.030[c]	26.8 (4.9)	27.9 (5.9)	0.267[c]
Median duration symptoms (months, IQ range)	4.0 (7.0)	6.0 (22.0)	0.078[c]	4.0 (10.0)	6.0 (22.0)	0.162[c]
Median wrist circumference (cm, IQ range)	16.0 (0.7)	17.0 (1.6)	<0.001[c]	16.0 (0.7)	17.0 (1.6)	<0.001[c]
Minimum-maximum range wrist circumference in cm	14.0–16.7	14.0–20.0		14.4–18.0	14.0–22.0	
Median CSA (mm^2, IQ range)	10.3 (0.6)	11.7 (4.4)	0.281[c]	11.0 (1.7)	11.3 (4.8)	0.618[c]

[a] *Pearson Chi-Square test.*
[b] *Unpaired t-test.*
[c] *Mann–Whitney U-Test.*

the other two cut-off methods, a statistically significantly smaller wrist circumference was noticed, as is shown in **Table 6**. These results may point out that a cut-off value based on the wrist-circumference adds the most value in people with a smaller wrist circumference when compared with the other analysed US parameters in this study.

The lack of a gold standard is an important problem in the diagnosis of CTS and this complicates research regarding CTS (15). In literature, clinical signs and symptoms, NCS and (surgical) outcome are used as reference standards (15). CTS is a clinical diagnosis and without signs and symptoms an individual cannot be diagnosed with CTS but can have abnormal US/NCS outcomes. False positives as well as false negatives are therefore present in groups of patients with clinically defined CTS but also in groups of patients with e.g., abnormal US and/or abnormal NCS outcomes. A positive effect of surgical treatment is another possible reference standard. However, even sham operations could have a positive (placebo) effect. Furthermore, in a previous study we showed that patients with clinically defined CTS and normal NCS noted a significant reduction of complaints after carpal tunnel release (16). For the examined wrists in this study the same problem of false positives and false negatives exists and the less strict inclusion criteria are probably the cause of the low sensitivity of US in this study.

Several studies suggest that increased nerve vascularisation in the median nerve, as evaluated by colour Doppler sonography, is associated with (severity of) carpal tunnel syndrome (10). However, one study showed an increased median nerve vascularisation prevalence of 36% in 60 healthy individuals (14). Surprisingly, in this study, median nerve vascularisation was only present in 10.9% of the participants and in the Radboudumc absent in all participants. This is not in line with literature where increased nerve vascularisation is reported in 41–95% of CTS patients (10, 17). This may be explained by differences in techniques while performing US (e.g., manual compression) or differences in settings and/or Doppler sensitivity between US devices. A cut-off value based on nerve median nerve vascularisation seemed to be the least favourable cut-off compared with the other investigated parameters in this study.

There are several limitations to this study. Firstly, it is important to bear in mind the possible bias in our data caused by our more liberal inclusion criteria for investigating carpal tunnel syndrome. We did not mean to change the clinical criteria for CTS in any way, but our goal was to investigate a patient population, as may be encountered in daily outpatient clinical practice, as described in the introduction. Secondly, the neurophysiology technicians performing the US were not blinded. The technicians may have expected to find enlargement of the median nerve, however, compared to an earlier Dutch study the median CSA of the median nerve at wrist level seems to be lower in this study, especially in the Zuyderland hospital (13). Concerning ultrasonography, we did not measure intra- or interobserver variability of the measurements. Earlier studies reported good agreement of CSA measurements of the median nerve (18, 19) but we cannot fully exclude interobserver variability due to variation in outlining the nerve contour (20), particularly because of the significantly smaller CSA found in the Zuyderland hospital. Only in the Radboudumc the distal carpal tunnel was visualised, in the other hospitals the median nerve was visualised only in the proximal and middle part of the carpal tunnel. However, in this study the percentage of abnormal US in the Canisius-Wilhelmina hospital and the Radboudumc were comparable. Furthermore, we cannot exclude slight differences in interhospital interpretations concerning the inclusion criteria.

In conclusion, a WDC for the CSA of the median nerve at the wrist appears to have a higher sensitivity than a FC or IFC in CTS patients in clinical practice, who do not always fulfil more rigid clinical criteria for the clinical diagnosis CTS. A cut-off for the CSA of the median nerve based on the wrist-circumference may at present be the most powerful approach in patients with a smaller wrist circumference.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Commissie Mensgebonden Onderzoek regio Arnhem-Nijmegen. The ethics committee waived the requirement of written informed consent for participation.

AUTHOR CONTRIBUTIONS

TO, FD, JM, and WV: conception and design of study. TO, RBe, and JW: acquisition of data. TO, FD, RBe, JW, RBa, JM, and WV: analysis and/or interpretation of data. TO: drafting of manuscript. TO, FD, RBe, JW, RBa, JM, and WV: revising manuscript. All authors read and approved the final version.

ACKNOWLEDGMENTS

We like to thank all participants of this study. Furthermore, we thank our electrodiagnostic technicians and physician assistants for their contributions to this study.

REFERENCES

1. Atroshi I, Gummesson C, Johnsson R, Ornstein E, Ranstam J, Rosén I. Prevalence of carpal tunnel syndrome in a general population. *J Am Med Assoc.* (1999) 282:153–8. doi: 10.1001/jama.282.2.153
2. De Krom MCTFM, Knipschild PG, Kester ADM, Thijs CT, Boekkooi PF, Spaans F. Carpal tunnel syndrome: prevalence in the general population. *J Clin Epidemiol.* (1992) 45:373–6. doi: 10.1016/0895-4356(92)90038-O
3. Katz JN, Simmons BP. Clinical practice. Carpal tunnel syndrome. *N Engl J Med.* (2002) 346:1807–12. doi: 10.1056/NEJMcp013018

4. Westerman D, Kerkhoff H, Visser GH, Kleyweg RP. Interobserver agreement in case history evaluation in carpal tunnel syndrome. *J Clin Neuromuscul Dis.* (2012) 13:196–200. doi: 10.1097/CND.0b013e31824619ad
5. Claes F, Verhagen WIM, Meulstee J. Current practice in the use of nerve conduction studies in carpal tunnel syndrome by surgeons in the netherlands. *J Hand Surg Eur Vol.* (2007) 32:663–7. doi: 10.1016/J.JHSE.2007.09.007
6. Nederlandse Vereniging voor Neurologie. Richtlijn carpaletunnelsyndroom (2017). Available online at: https://richtlijnendatabase.nl/richtlijn/carpaletunnelsyndroom_cts/startpagina_-_carpaletunnelsyndroom_cts.html
7. Tai TW, Wu CY, Su FC, Chern TC, Jou IM. Ultrasonography for diagnosing carpal tunnel syndrome: a meta-analysis of diagnostic test accuracy. *Ultrasound Med Biol.* (2012) 38:1121–8. doi: 10.1016/j.ultrasmedbio.2012.02.026
8. Buchberger W, Judmaier W, Birbamer G, Lener M, Schmidauer C. Carpal tunnel syndrome: diagnosis with high-resolution sonography. *Am J Roentgenol.* (1992) 159:793–8. doi: 10.2214/ajr.159.4.1529845
9. Ulaşli AM, Duymuş M, Nacir B, Rana Erdem H, Koşar U. Reasons for using swelling ratio in sonographic diagnosis of carpal tunnel syndrome and a reliable method for its calculation. *Muscle and Nerve.* (2013) 47:396–402. doi: 10.1002/mus.23528
10. Vanderschueren GAKB, Meys VEWH, Beekman R. Doppler sonography for the diagnosis of carpal tunnel syndrome: a critical review. *Muscle and Nerve.* (2014) 50:159–63. doi: 10.1002/mus.24241
11. Olde Dubbelink TBG, De Kleermaeker FGCM, Meulstee J, Bartels RHMA, Claes F, Verhagen WIM. Augmented diagnostic accuracy of ultrasonography for diagnosing carpal tunnel syndrome using an optimised wrist circumference-dependent cross-sectional area equation. *Front Neurol.* (2020) 11:577052. doi: 10.3389/fneur.2020.577052
12. Claes F, Kasius KM, Meulstee J, Grotenhuis JA, Verhagen WIM. Treatment outcome in carpal tunnel syndrome: does distribution of sensory symptoms matter? *J Neurol Sci.* (2014) 344:143–8. doi: 10.1016/j.jns.2014.06.044
13. Visser LH, Smidt MH, Lee ML. High-resolution sonography versus EMG in the diagnosis of carpal tunnel syndrome. *J Neurol Neurosurg Psychiatry.* (2008) 79:63–7. doi: 10.1136/jnnp.2007.115337
14. Houben MPWA, Maars M, Beekman R. High prevalence of increased nerve vascularization in healthy individuals. *Muscle and Nerve.* (2015) 51:938–9. doi: 10.1002/mus.24655
15. Sonoo M, Menkes DL, Bland JDP, Burke D. Nerve conduction studies and EMG in carpal tunnel syndrome: do they add value? *Clin Neurophysiol Pract.* (2018) 3:78–88. doi: 10.1016/j.cnp.2018.02.005
16. De Kleermaeker FGCM, Meulstee J, Claes F, Kasius KM, Verhagen WIM. Treatment outcome in patients with clinically defined carpal tunnel syndrome but normal electrodiagnostic test results: a randomized controlled trial. *J Neurol.* (2017) 264:2394–400. doi: 10.1007/s00415-017-8637-2
17. Kutlar N, Bayrak AO, Bayrak IK, Canbaz S, Türker H. Diagnosing carpal tunnel syndrome with Doppler ultrasonography: a comparison of ultrasonographic measurements and electrophysiological severity. *Neurol Res.* (2017) 39:126–32. doi: 10.1080/01616412.2016.1275455
18. Claes F, Meulstee J, Claessen-Oude Luttikhuis TTM, Huygen PLM, Verhagen WIM. Usefulness of additional measurements of the median nerve with ultrasonography. *Neurol Sci.* (2010) 31:721–5. doi: 10.1007/s10072-010-0258-9
19. Nakamichi KI, Tachibana S. Enlarged median nerve in idiopathic carpal tunnel syndrome. *Muscle Nerve.* (2000) 23:1713–8. doi: 10.1002/1097-4598(200011)23:11<1713::AID-MUS7>3.0.CO;2-G
20. Coraci D, Giovannini S, Loreti C, Ruggeri F, Padua L. The hyperchoic rim of the normal nerve in ultrasound: how significant is it? *Neurol Sci.* (2020) 41:2985–7. doi: 10.1007/s10072-020-04405-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 15 February 2021
doi: 10.3389/fneur.2021.608549

Diffusion Tensor Imaging of Skeletal Muscle Contraction Using Oscillating Gradient Spin Echo

Valentina Mazzoli, Kevin Moulin, Feliks Kogan, Brian A. Hargreaves and Garry E. Gold*

Department of Radiology, Stanford University, Stanford, CA, United States

Diffusion tensor imaging (DTI) measures water diffusion in skeletal muscle tissue and allows for muscle assessment in a broad range of neuromuscular diseases. However, current DTI measurements, typically performed using pulsed gradient spin echo (PGSE) diffusion encoding, are limited to the assessment of non-contracted musculature, therefore providing limited insight into muscle contraction mechanisms and contraction abnormalities. In this study, we propose the use of an oscillating gradient spin echo (OGSE) diffusion encoding strategy for DTI measurements to mitigate the effect of signal voids in contracted muscle and to obtain reliable diffusivity values. Two OGSE sequences with encoding frequencies of 25 and 50 Hz were tested in the lower leg of five healthy volunteers with relaxed musculature and during active dorsiflexion and plantarflexion, and compared with a conventional PGSE approach. A significant reduction of areas of signal voids using OGSE compared with PGSE was observed in the tibialis anterior for the scans obtained in active dorsiflexion and in the soleus during active plantarflexion. The use of PGSE sequences led to unrealistically elevated axial diffusivity values in the tibialis anterior during dorsiflexion and in the soleus during plantarflexion, while the corresponding values obtained using the OGSE sequences were significantly reduced. Similar findings were seen for radial diffusivity, with significantly higher diffusivity measured in plantarflexion in the soleus muscle using the PGSE sequence. Our preliminary results indicate that DTI with OGSE diffusion encoding is feasible in human musculature and allows to quantitatively assess diffusion properties in actively contracting skeletal muscle. OGSE holds great potential to assess microstructural changes occurring in the skeletal muscle during contraction, and for non-invasive assessment of contraction abnormalities in patients with muscle diseases.

Keywords: MRI, DTI, OGSE, diffusion MRI, oscillating gradients, muscle contraction

OPEN ACCESS

Edited by:
Francesco Santini, University of Basel, Switzerland

Reviewed by:
Eric Edward Sigmund, New York University, United States
Petros Martirosian, Tübingen University Hospital, Germany

****Correspondence:***
Valentina Mazzoli
vmazzoli@stanford.edu

Specialty section:
This article was submitted to Neuromuscular Diseases, a section of the journal Frontiers in Neurology

Received: *21 September 2020*
Accepted: *08 January 2021*
Published: *15 February 2021*

Citation:
Mazzoli V, Moulin K, Kogan F, Hargreaves BA and Gold GE (2021) Diffusion Tensor Imaging of Skeletal Muscle Contraction Using Oscillating Gradient Spin Echo. Front. Neurol. 12:608549. doi: 10.3389/fneur.2021.608549

INTRODUCTION

Diffusion tensor imaging (DTI) is an MRI-based technique that allows to measure the anisotropic diffusion of water molecules in muscle tissue. DTI can non-invasively provide *in vivo* information on tissue architecture and microstructure, either normal or in a diseased state. In a typical DTI experiment, diffusion is probed along multiple directions using diffusion encoding gradients. The application of diffusion encoding gradients results in a signal attenuation, which across several directions can be geometrically interpreted by a rank 2 tensor. This tensor can be

diagonalized to derive the principal directions of diffusion. In skeletal muscle, the highest diffusivity is observed along the axis of the fiber [axial diffusivity (AD)] and the lowest in the fiber cross section [radial diffusivity (RD)]. Due to its exquisite sensitivity to tissue microstructure, DTI is becoming an increasing popular tool to assess skeletal muscle status in a wide range of diseases (1) and muscle injuries (2).

Muscle contraction involves muscle fiber shortening and increase in cross-sectional area (CSA). The ability of accurately measuring these structural changes during muscle contraction *in vivo* has the potential to elucidate contraction abnormalities that are not apparent in relaxed musculature. Extensive simulation work showed the high sensitivity of skeletal muscle DTI to cellular size (3), with higher diffusion coefficients and lower fractional anisotropy (FA) associated with increasing fiber CSA. Due to its high cellular size sensitivity, DTI is a very promising tool to assess *in vivo* and non-invasively skeletal muscle microstructure and microstructural changes due to contraction in human musculature. Previous work has shown the high sensitivity of DTI to transient structural changes induced by passive shortening and lengthening (4–6) with increased radial diffusivity associated with passive shortening and increase in the fascicle CSA. Active muscle contraction induces shortening of muscle fibers, associated with increased fiber CSA. DTI is therefore an attractive tool to non-invasively assess muscle contraction mechanism by measuring changes in radial diffusivity during active contraction with respect to relaxed musculature.

DTI has been extensively applied to probe skeletal muscle microstructure in healthy subjects and patients. However, most of this work has been performed for static, non-contracting muscle, mostly due to the presence of signal voids in diffusion-weighted images acquired during muscle contraction (7, 8). These areas of signal void are caused by incoherent motion within the contracting muscle tissue (9), which results in rapid signal dephasing in contracting muscle and preclude further analyses and extraction of microstructural information from the DTI measurement. This need for static acquisition largely prohibits the characterization of fundamental functional aspects of the skeletal muscle, such as microstructural changes due to contraction and contraction-induced diffusion response. Recent work has shown that the presence and extent of signal voids occurring during external muscle stimulation in diffusion MRI can provide information on motor units (10). However, these methods do not allow to characterize the microstructural features of skeletal muscle.

Diffusion behavior in skeletal muscle has long been shown to be highly time-dependent (11), with diffusion coefficients measured using DTI that strongly depends on the time allowed for diffusing water molecules to probe the local environment (the so-called "diffusion dime"). This time-dependent behavior has also been combined with advanced mathematical modeling to derive cell size (12). For increasing diffusion times, the water molecules will interact with more barriers, and the apparent diffusion coefficient (ADC) will decrease, eventually reaching an asymptotic lower value (13, 14). Skeletal muscle is a highly hierarchical structure, and it is therefore possible to study different levels of tissue organization by tuning the diffusion time accordingly.

DTI in skeletal muscle is typically performed using pulsed gradient spin echo (PGSE) diffusion encoding, which has an intermediate diffusion time (about 20-30 ms), resulting in an inability to provide information at the level of the individual muscle fibers, typically on the order of 30–60μm. In practice, PGSE is also extremely sensitive to bulk motion caused by voluntary or involuntary muscle contraction, which leads to unwanted signal dropouts in the acquired images. Diffusion methods with very long diffusion times such as stimulated echoes are also routinely used to investigate muscle fiber size (15) but suffer from the same issue as the PGSE approach. OGSE has been proposed in the brain imaging field as an efficient method to reduce the diffusion time of the experiment while maintaining a sufficient amount of diffusion encoding (16, 17). By design, cosine or trapezoid-cosine OGSE waveforms also provide full motion compensation ($M_0 = M_1 = M_2 = 0$), which could be particularly useful to compensate for bulk tissue motion induced by muscle contraction. Additionally, its short diffusion times could allow to investigate ultrastructural features that are currently not accesible using diffusion encoding schemes with longer diffusion times. OGSE could also allow to study microstructural and ultrastructural changes in actively contracted muscles. The OGSE method, although promising, is mostly confined to neurological (16, 18) and cancer (19) application and has never been applied to the skeletal muscle to date.

Therefore, the aim of this work is to explore OGSE for evaluation of the human skeletal muscle on a clinical 3T MRI scanner and to exploit its inherent motion compensation to measure muscle microstructure during active muscle contraction.

METHODS

Subjects

Five healthy volunteers (four females and one male) were recruited for lower-leg MRI imaging. None of the subjects had a history of muscle disease, and they were asked to refrain from any strenuous physical activities the day before the scan. This study was approved by the University Institutional Review Board. We received informed consent from all subjects prior to the study, according to our institution's regulations.

Data Acquisition

MR imaging was performed with a 3T MRI scanner (SIGNA 750w Premier, GE Healthcare) with maximal nominal gradient strength = 80 mT/m and max nominal slew rate = 120 mT/m/s. A receive-only medium-size 16-channel flexible coil array (NeoCoil, Pewaukee, WI, USA) wrapped around the left calf was used for signal reception. All subjects were placed supine, feet first in the scanner, with both legs in the full extended position. The scan protocol consisted of a series of DTI scans acquired with relaxed musculature and during foot dorsiflexion and plantarflexion. Sandbags on both sides of the left leg were used to minimize displacement of the subject between

acquisitions. DTI scans were performed in three configurations: (1) at rest, (2) with the subjects actively contracting their muscle to maintain a plantarflexion position, while pushing their foot against a rigid support, and (3) with the subject actively contracting their muscles to achieve the maximum level of dorsiflexion (**Figure 1**). Three different DTI sequences (PGSE, OGSE 25 Hz and OGSE 50 Hz) were collected for each position (no contraction, dorsiflexion, and plantarflexion), resulting in nine DTI acquisitions per subject. The order of positions and DTI acquisitions was randomized to reduce the effect of perfusion and muscle fatigues on the DTI results. For the OGSE acquisition, cosine trapezoidal gradient waveforms (20) with 50 and 25 Hz, corresponding to $N = 2$ and $N = 1$ oscillations, were implemented. This resulted in a diffusion time of 7.5 and 4.1 ms for the OGSE 25 Hz and OGSE 50 Hz sequences, respectively (16). For both OGSE frequencies, the encoding waveform duration before and after the refocusing pulse was 40 ms long (**Figure 2**). The PGSE sequence was designed to match the timing of the OGSE sequence and was thus composed of a single monopolar trapezoid gradient with 40-ms duration. For each DTI acquisition, the maximum gradient strength used was adjusted in order to achieve the same maximum b-value of 180 s/mm^2. For the three DTI acquisitions (PGSE, OGSE 25 Hz, and OGSE 50 Hz), diffusion was encoded along 15 non-collinear diffusion encoding directions, and three non-diffusion-weighted volumes were acquired. Two averages were acquired for the diffusion-weighted volume, for a total of 33 scanned volumes per sequence. Other common scan parameters were repetition time/echo time (TR/TE) = 2,800/94 ms, field of view (FOV) = 160 × 160 mm^2, 10 slices, voxel size = 2.7 × 2.7 × 10 mm^3. Spectral spatial water excitation was used for fat suppression. The total scan time for each DTI dataset was 1 min 30 s.

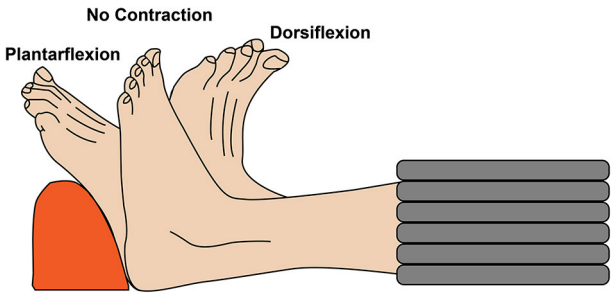


FIGURE 1 | Schematic representation of the experimental setup. The left lower leg of the volunteers was scanned with relaxed musculature and during active foot dorsiflexion and plantarflexion.

Image Processing

In order to estimate the signal-to-noise ratio (SNR), we calculated the standard deviation of the noise (σ) from two identically acquired non-diffusion-weighted and diffusion-weighted images ($b = 0$ mm^2/s and $b = 180$ mm^2/s) for all the scans acquired in neutral position. SNR was defined as the mean of the signal over all muscles in the lower leg divided by σ. All images were visually inspected, and the areas of signal voids in the anterior and posterior compartments of the lower leg in each DTI scan were counted in a mid-calf slice (Slice 4) in 30 images (15 diffusion encoding directions * 2 averages). Areas of signal voids were classified as area of four or more contiguous black pixels in the muscle of interest.

All DTI scans were denoised using a principal component analysis (PCA) denoising algorithm (21). The non-diffusion-weighted scan acquired using the PGSE sequence with the lower leg in relaxed position was used as an anatomical reference. For each subject, all other DTI datasets were non-rigidly registered to the reference scan, in order to take into account the change in shape of the muscle due to active contraction. All image registrations were performed using Elastix (22), with a b-spline algorithm with the metric Advanced Mattes Mutual information = and a b-spline interpolation order of three. The number of resolutions for the registration was set to two. All image registrations was performed using Elastix (22). The reference image was also used to manually delineate the tibialis anterior and soleus muscles. After registration, the diffusion images per acquisition were fitted into a DTI model using a WLLS algorithm

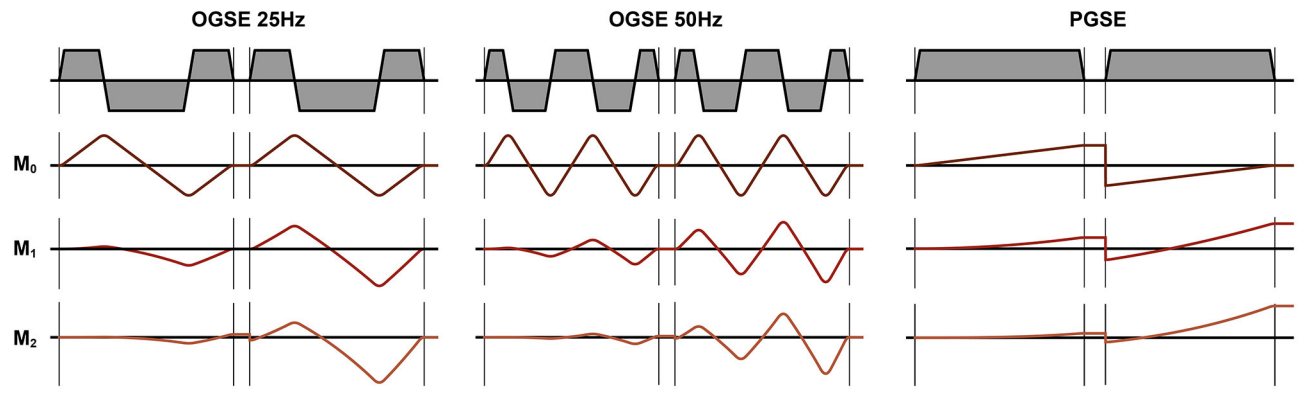


FIGURE 2 | Diffusion encoding gradients for oscillating gradient spin echo (OGSE) 25 Hz, OGSE 50 Hz, and pulsed gradient spin echo (PGSE). The OGSE waveforms are motion compensated ($M_0 = M_1 = M_2 = 0$), while the PGSE is not motion compensated ($M_0 = 0$, $M_1 \neq 0$, $M_2 \neq 0$).

(23) with outlier rejections (24). The first eigenvalue of the DTI tensor correspond to the AD ($=\lambda_1$) and the mean of the second and third eigenvalues to the RD [$=(\lambda_2 + \lambda_3)/2$]. All image processing and tensor fitting was performed using QMRITools (25) (https://mfroeling.github.io/QMRITools).

Statistical Analysis

Differences in number of signal voids, AD, RD, and FA, were tested using a mixed-model ANOVA, accounting for subject, DTI sequence (PGSE, OGSE 25 Hz, and OGSE 50 Hz), and leg position (no contraction, active dorsiflexion, and active plantarflexion). *Post-hoc* Tukey test was used to examine individual relationships. All statistical analyses were performed using SPSS. Additionally, a one-way ANOVA with a Fisher correction was used to test the effect of contraction status for the diffusion parameters obtained with OGSE 25 Hz and OGSE 50 Hz sequences. All tests were performed two-sided with significance set at $p < 0.05$.

RESULTS

All results were visually inspected and were considered of sufficient quality to continue with further analyses. The mean SNR for the non-diffusion-weighted scan was 32 ± 2, 28 ± 2, and 28 ± 2 for OGSE 25 Hz, OGSE 50 Hz, and PGSE, respectively. The mean SNR for the diffusion-weighted scans was 16 ± 4, 16 ± 1, and 17 ± 1 for OGSE 25 Hz, OGSE 50 Hz, and PGSE, respectively.

Figure 3 summarizes the results of visual counting of areas of extended signal voids in diffusion-weighted volumes at different contractions and using different diffusion encoding waveforms. For the images acquired with relaxed musculature, areas of signal voids were present for each volunteer using the PGSE sequence. These areas of signal voids, likely originating by involuntary muscle twitch (7), were localized in the posterior compartment of the leg and were not present when encoding diffusion using OGSE waveforms. No signal voids were observed in relaxed musculature in the tibialis anterior muscle with any of the waveforms. Images acquired using PGSE diffusion encoding during active plantarflexion had significantly more areas of signal voids in the soleus than using OGSE 25 Hz and OGSE 50 Hz ($p < 0.0001$). These artifacts were present in every subject. OGSE waveforms resulted in a significantly reduced number of signal voids (observed in three subjects and one subject for the OGSE 25 Hz and OGSE 50 Hz, respectively). Similarly, areas of signal voids were observed in the tibialis anterior muscle for every volunteer during active dorsiflexion but were significantly reduced when using OGSE ($p < 0.0001$ for both OGSE 25 Hz and OGSE 50 Hz) for the same type of contraction.

Representative images of the calf of a female volunteer showing images acquired using OGSE 25 Hz, OGSE 50 Hz, and PGSE are shown in **Figure 4**. PGSE scans acquired during active muscle contraction showed clear areas of signal voids localized in the anterior compartment of the lower leg (for the dorsiflexed foot position) and in the posterior compartment (for the plantarflexed foot position). The presence of these areas was dramatically reduced in the scans acquired using OGSE 25 Hz and OGSE 50 Hz, with either relaxed or contracted musculature.

Representative RD maps for one volunteer acquired with the three different sequences and three different contraction status are shown in **Figure 5**. The areas of signal voids present in the diffusion-weighted images acquired with PGSE during active dorsiflexion and plantarflexion resulted in abnormally elevated values of AD and RD in the soleus and tibialis anterior muscles, respectively (**Figures 6**, **7**).

The results of a mixed-model ANOVA showed significantly higher AD values in the tibialis anterior muscle during active dorsiflexion than in plantarflexion ($p = 0.021$) and no-contraction positions ($p = 0.031$). These differences were driven by the abnormally elevated AD values measured in dorsiflexion using PGSE (3.05 ± 0.51 mm^2/s) compared with OGSE 25 Hz (2.37 ± 0.51 mm^2/s) and OGSE 50 Hz (2.35 ± 0.75 mm^2/s). Elevated AD values in the tibialis anterior were observed for the PGSE sequence compared with OGSE 50 Hz ($p = 0.024$). Abnormally elevated AD values were observed in the soleus during plantarflexion using the PGSE sequence (3.40 ± 1.05 mm^2/s), but these were in a normal range when using the OGSE sequences (2.45 ± 0.41 mm^2/s for OGSE 25 Hz and 2.37 ± 0.28 mm^2/s for OGSE 50 Hz). Similar trends of elevated AD were also observed in the soleus when using the PGSE sequence (PGSE vs. OGSE 25 Hz, $p = 0.053$), with higher AD values during plantarflexion compared with dorsiflexion ($p = 0.026$) and no contraction ($p = 0.001$).

The OGSE sequences detected significantly higher AD in the soleus during active plantarflexion compared with no-contraction position ($p = 0.013$), but no significant differences were observed for the other positions ($p > 0.081$). No differences in AD were detected between positions for the tibialis anterior ($p > 0.117$), but a trend of increasing AD in dorsiflexion was observed.

We observed higher RD values in the tibialis anterior muscle when using PGSE during active dorsiflexion (1.68 ± 0.18 mm^2/s for PGSE vs. 1.52 ± 0.38 mm^2/s for OGSE 25 Hz and 1.56 ± 0.75 mm^2/s), although the interaction of sequence and position was not significant ($p = 0.928$). The RD values in the soleus, similarly to AD, were significantly higher in plantarflexion compared with no-contraction position ($p = 0.010$), but no significant difference was detected between plantarflexion and dorsiflexion positions ($p = 0.579$). These differences were primarily driven by the higher RD values observed in plantarflexion using the PGSE sequence (1.88 ± 0.34 mm^2/s) compared with OGSE 25 Hz (1.54 ± 0.37 mm^2/s) and OGSE 50 Hz (1.53 ± 0.25 mm^2/s).

When using an OGSE approach, RD in the soleus during no contraction was significantly lower than in plantarflexion ($p = 0.028$) and dorsiflexion ($p = 0.026$). No differences in RD were observed for the tibialis anterior between different leg positions ($p > 0.325$), although a trend of increasing RD in dorsiflexion was observed.

FA values in the soleus and tibialis anterior muscles, averaged over all volunteers, are reported in **Figure 8**. In the tibialis anterior, the FA values were significantly lower in dorsiflexion compared with plantarflexion ($p = 0.004$) and no-contraction

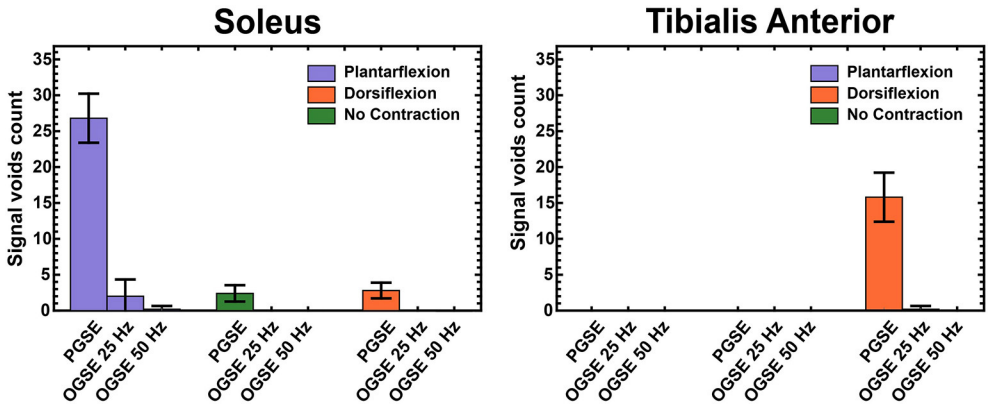


FIGURE 3 | Counts of areas with extended signal voids in Slice 4 (mid-calf) in the soleus and tibialis anterior muscles, averaged over all volunteers.

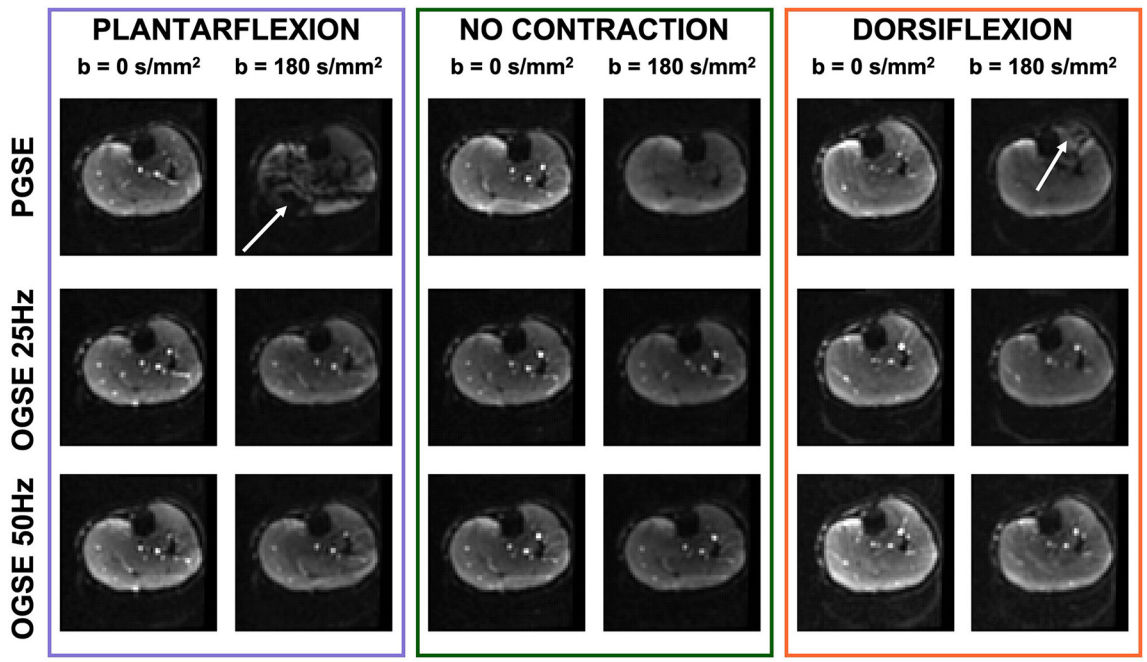


FIGURE 4 | Non-diffusion-weighted and diffusion-weighted scans of one volunteer acquired during active plantarflexion (purple), with relaxed musculature (no contraction, green), and during active dorsiflexion (orange). The pulsed gradient spin echo (PGSE) scans show localized areas of signal voids (white arrow) originating by spin dephasing during muscle activation. Note the bright signal from blood vessels in the diffusion-weighted scans obtained with oscillating gradient spin echo (OGSE), indicating motion compensation.

position ($p = 0.003$). An interaction effect was observed for FA in the soleus ($p = 0.004$), but no separate effect of position ($p = 0.303$) or sequence ($p = 0.341$). No significant differences were observed between OGSE 25 Hz and OGSE 50 Hz sequences for any of the diffusion quantitative parameters ($p > 0.1$).

No significant differences for different foot positions were observed for FA in either soleus ($p > 0.090$) or tibialis anterior ($p > 0.102$) muscles when using OGSE.

DISCUSSION

This study applied DTI with OGSE diffusion encoding in the musculature of the lower leg to characterize diffusion behavior in actively contracting skeletal muscle for the first time. More than half of the diffusion-weighted volumes acquired using the PGSE sequence during active muscle contraction showed extended areas of signal voids that led to an unrealistic overestimation of diffusion parameters.

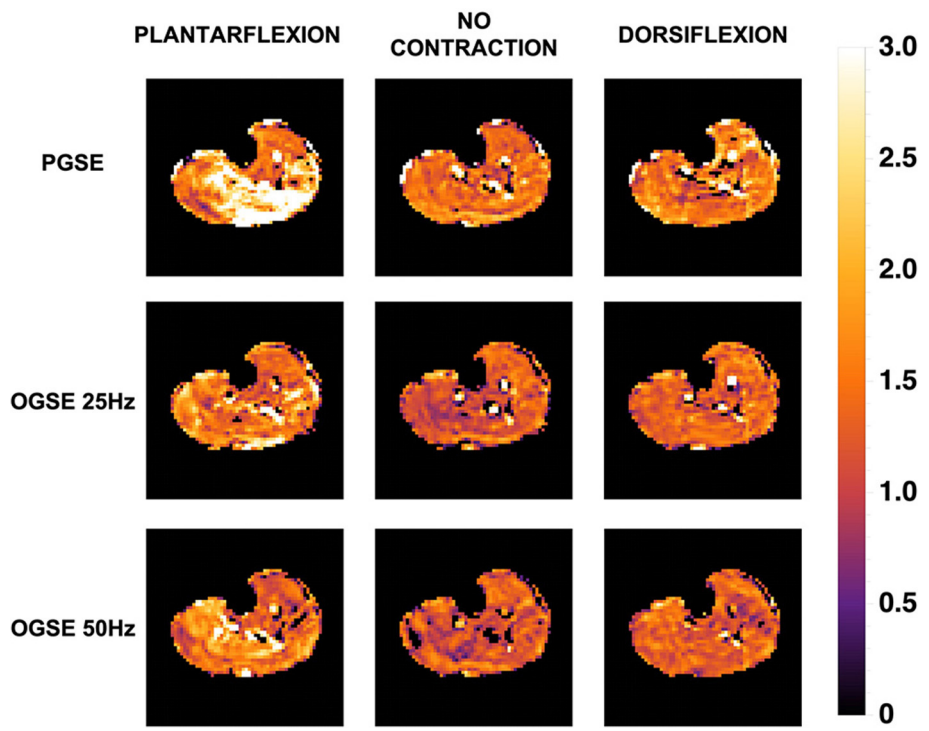


FIGURE 5 | Radial diffusivity (RD) maps for one female volunteer. Abnormally elevated diffusion values are observed in the posterior compartment of the lower leg during plantarflexion and in the anterior compartment during dorsiflexion when using a pulsed gradient spin echo (PGSE) sequence.

These artifacts were significantly reduced when using the OGSE sequences and led to physically plausible diffusion values, indicating that trapezoid-cosine OGSE waveforms, due to their motion compensation design, might be promising to investigate muscle microstructure during active muscle contraction.

Areas of signal voids in healthy resting muscle at moderate b-values and associated with a spontaneous incoherent mechanical activity result in local signal dephasing when using stimulated echoes (9) and need to be discarded prior to further analyses and DTI fitting (26). Steidle et al. (9) reported clear and extended areas of signal in resting musculature or the lower leg at b-values as low as 100 s/mm^2 and a higher occurrence of signal voids in the posterior compartment compared with the anterior compartment, similarly to what was observed in our study. These signal voids, observed in resting musculature, have been associated with incoherent motion of muscle fibers due to a spontaneous muscle activity (7–9). Electrical muscle stimulation of calf muscle has been shown to increase the amount and spatial extent of signal voids in diffusion-weighted scans acquired at low b-values along a single direction, suggesting the sensitivity of diffusion MRI to size and shape of motor units (10, 27). While these results obtained during external stimulation clearly demonstrate the potential of diffusion MRI to study contraction mechanisms, being able to assess muscle microstructure during active contraction would increase the clinical applicability of the method.

The use of PGSE waveforms led to unrealistically and non-physical AD values when performing DTI experiments in actively contracting muscle (soleus during plantarflexion and tibialis anterior during dorsiflexion), while no differences were observed for between the two OGSE sequences. OGSE showed increased RD in the soleus during plantarflexion compared with neutral position, possibly indicating increased muscle CSA due to active contraction, similar to previous findings in passively shortened muscle (4–6, 28, 29). Since increased CSA is expected for increasing applied force, future investigations will focus on the effect of force on measured RD. However, while AD in passively contracted muscles has been shown to be largely unaffected, interestingly in this study, we observed an increase in AD in the soleus during plantarflexion. Muscle contraction involves shortening of the sarcomeres. However, since the shortest diffusion time in our study would be associated with distances larger than the typical sarcomere size, the observed differenced in AD are likely caused by other restrictions. Further research is warranted to elucidate the cause for the observed changes in AD upon muscle contraction.

The motion of actively contracting tissue is expected to be somewhat coherent, and our results clearly indicate that trapezoid-cosine OGSE waveforms, being motion compensated, are able to fully refocus spin dephasing due to coherent motion.

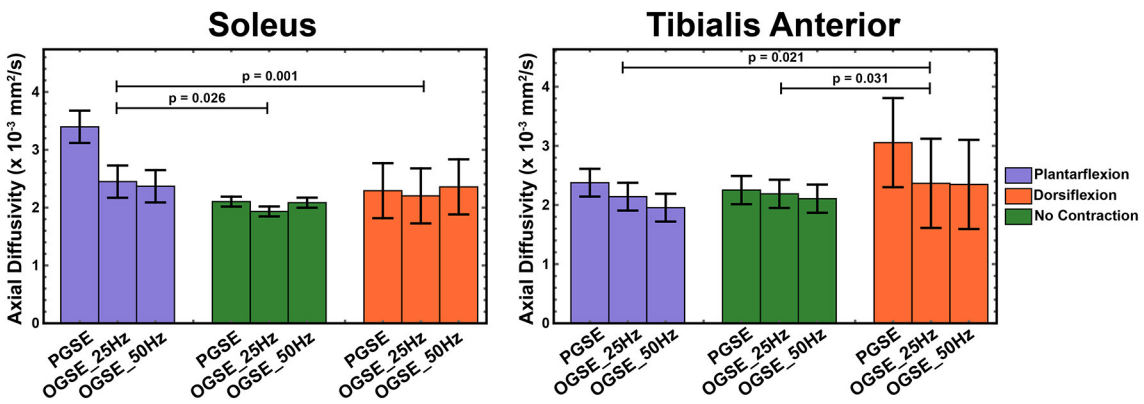


FIGURE 6 | Axial diffusivity (AD) measured during active plantarflexion (purple), with relaxed musculature (green), and during active dorsiflexion (orange) using three different diffusion encoding sequences [pulsed gradient spin echo (PGSE), oscillating gradient spin echo (OGSE) 25 Hz, and OGSE 50 Hz]. Elevated AD values are present in the soleus during active plantarflexion and in the tibialis anterior during active dorsiflexion.

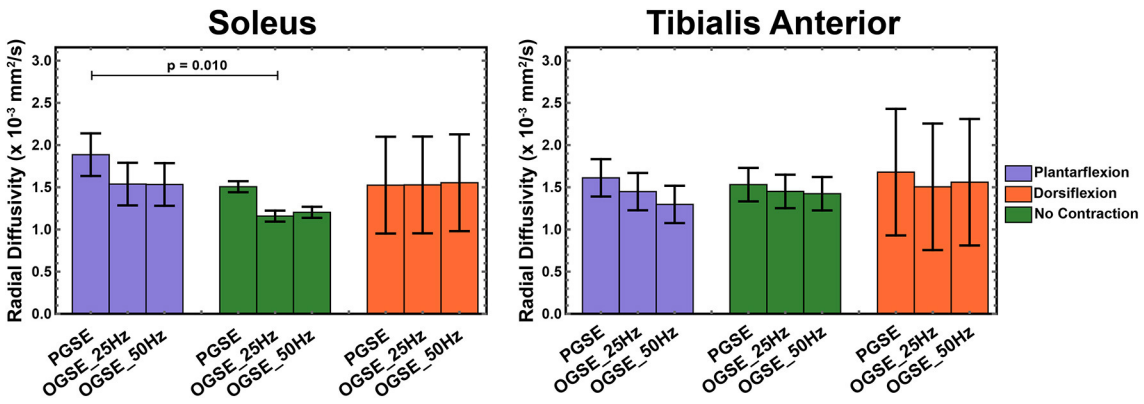


FIGURE 7 | Radial diffusivity (RD) measured during active plantarflexion (purple) and active dorsiflexion (orange) and with relaxed musculature (green) using three different diffusion encoding sequences [pulsed gradient spin echo (PGSE), oscillating gradient spin echo (OGSE) 25 Hz, and OGSE 50 Hz]. Elevated RD values are present in the soleus during active plantarflexion.

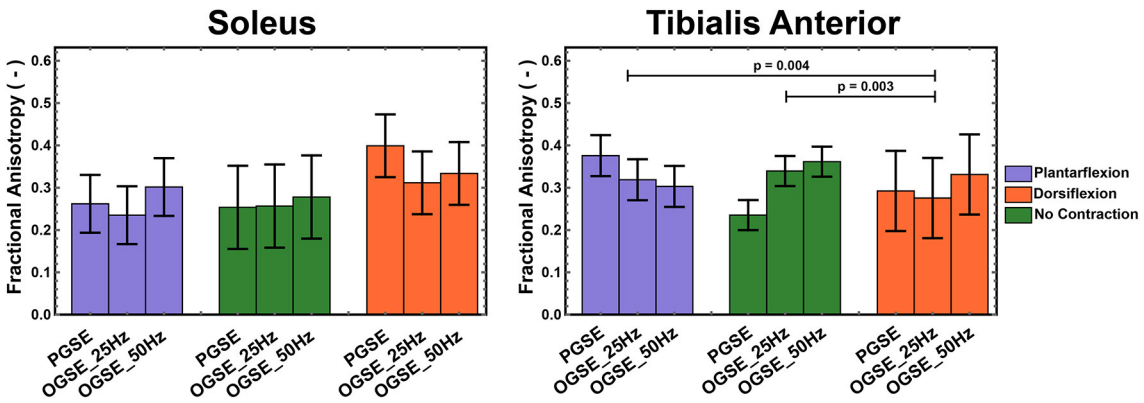


FIGURE 8 | Fractional anisotropy (FA) measured during active plantarflexion (purple) and active dorsiflexion (orange) and with relaxed musculature (green) using three different diffusion encoding sequences [pulsed gradient spin echo (PGSE), oscillating gradient spin echo (OGSE) 25 Hz, and OGSE 50 Hz]. Significantly lower FA values were measured in the tibialis anterior during active foot dorsiflexion.

While residual areas of signal voids are observed even when encoding diffusion using an OGSE approach, robust tensor fitting with outlier rejection offers robustness toward these artifacts and allows to obtain reliable diffusion estimation. While diffusion values for actively contracted skeletal muscle have not been previously reported, our AD and RD results obtained in relaxed muscle are in agreement with previously reported values obtained using conventional PGSE DTI in relaxed musculature (1). Mazzoli et al. performed a DTI experiment in relaxed calves using a PGSE approach and reported RD values of 1.55 ± 0.06 mm^2/s for the tibialis anterior and 1.51 ± 0.04 mm^2/s for the soleus and AD values of 2.32 ± 0.05 and 2.64 ± 0.07 mm^2/s (4). Similarly, Schlaffke et al. (30) reported 1.89 ± 0.08 and 2.14 ± 0.22 mm^2/s for AD and 1.35 ± 0.13 and 1.62 ± 0.21 mm^2/s for RD, for the tibialis anterior and soleus, respectively. Therefore, our results indicate that OGSE could be used to obtain a reliable diffusivity assessment in the lower leg. We should point out that the limited gradient strength of clinical MRI systems, together with peripheral nerve stimulation (PNS) constraints, resulted in relatively low b-value and long TE, which are suboptimal for DTI measurements in the skeletal muscle. However, while the OGSE sequence might not be optimal for routine DTI measurements, it could be highly beneficial to study actively contracted muscles or to image subjects who cannot keep their musculature fully relaxed during the full duration of the DTI experiment, as shown by its robustness to (in)voluntary muscle contraction.

Diffusion in the skeletal muscle is conventionally measured using a PGSE monopolar diffusion scheme with moderately short gradients and large gradient amplitude (28), resulting in moderate diffusion times. This monopolar approach however is not suitable if one wants to investigate time-dependent diffusion in the short time regime, due to its inherent inability to achieve sufficient diffusion encoding at short diffusion times. On the other hand, OGSE can achieve higher diffusion sensitivity for short diffusion times. Additionally, even though short gradients are less sensitive to motion compared with the longer PGSE gradients used in this study, they are still highly susceptible to signal dephasing due to contraction (31) and, therefore, unsuitable to obtain DTI information during muscle active muscle contraction. Our study suggests that fully compensated PGSE waveforms ($M_0 = M_1 = M_2 = 0$) could provide the same benefit as OGSE for skeletal muscle contraction imaging, although this hypothesis has to be tested. If time-dependent diffusion is not of interest, PGSE with first- and second-order motion compensation might provide a higher b-value than OGSE (32), which could be further improved by numerical optimization (33, 34).

DTI experiments using a monopolar PGSE waveform, which is not motion compensated, resulted in complete suppression of signal originating from flowing spins, as indicated by the dark blood vessels. On the other hand, when flow-compensated gradient waveforms, such as cosine OGSE, are used, coherent blood flow is no longer dephased, as indicated by the bright vessels in diffusion-encoded volumes, as previously observed (35). This could create an overestimation of diffusion values close to vessels due to partial volume effects when using compensated vs. uncompensated gradient waveforms. However, care was taken to avoid blood vessels during manual delineation of the tibialis anterior and soleus muscles, and our results show that RD and AD in resting musculature are higher for PGSE than for OGSE approaches even for non-contracting musculature. Therefore, partial volume effects with uncompensated blood flow are expected to have a limited effect on our results.

The measured AD and RD values did not show a clear time-dependent behavior in the soleus and tibialis anterior. Time-dependent diffusion in the cross section of skeletal muscle fibers has been investigated at longer diffusion times (up to 1 s) using stimulated echo approaches (13, 14), with lower diffusion coefficients measured for increasing diffusion times, likely indicating greater restriction effects of water molecules by cellular membranes. On the other hand, this study shows higher diffusivity values for PGSE, which has a longer diffusion time. Perfusion effects can lead to overestimation of diffusion coefficients (36), and the higher values of diffusivity observed in this study for the PGSE sequence compared with OGSE even in relaxed musculature could be explained by higher sensitivity to perfusion effects due to the lack of motion compensation (35). Therefore, based on this preliminary work, OGSE sequences with lower frequencies could be preferred, due to the possibility of achieving higher b-values for the same TE, which could further reduce the residual effect of perfusion on the measurements. The presence of a clear anisotropic diffusion behavior but no increase of diffusion values as a function of diffusion time (7.5 and 4.1 ms for OGSE 25 Hz and OGSE 50 Hz, respectively) could indicate a restricted diffusion behavior for all investigated sequences. This restricted diffusion behavior is likely not caused by the myofiber membranes, which are characterized by longer length scales but rather by ultrastructural restrictions. Further research is needed to elucidate the connection between the measured diffusion values and the underlying tissue structure.

This study has a number of limitations. First, this proof-of-concept study only included a limited number of subjects. Future research will focus on applying this method to study muscle contraction mechanisms in a larger number of healthy subjects and patients with neuromuscular diseases. Another potential limitation of our study is the relatively long TE of the acquisition, required to obtain a sufficient amount of diffusion encoding. This could have biased our analysis by increasing diffusion sensitivity to the extracellular space. Additionally, the limited gradient strength available on our clinical system, combined with compliance with PNS requirements, resulted in relatively low diffusion encoding strength, which might generate sensitivity to perfusion. Active muscle contraction could locally increase blood perfusion, which could lead to an overestimation of DTI parameters (36). Wu et al. (37) showed the dependency of diffusion contribution to the diffusion signal through the intravoxel incoherent motion (IVIM) model to be dependent on the average distance that blood can travel during the diffusion times, with smaller perfusion coefficients measured at shorter diffusion times. This work therefore suggests the potential of measurements at shorter diffusion times to minimize the effect of perfusion. Additionally, fully motion-compensated waveforms,

such as the trapezoid-cosine waveforms used in this study, were also shown to minimize the influence of perfusion in the liver in Moulin et al. (35), but our result could still be partially biased by residual perfusion contributions. Future work will focus on combining OGSE diffusion encoding with additional orthogonal gradients, as described by Wu et al. (37) for brain applications, to completely eliminate the influence of residual perfusion on the quantitative diffusion results. This promising method, However, we should mention that this promising method was implemented on a preclinical system with much higher gradient strength than currently achievable on a clinical system and will therefore require adaptations for applications in human skeletal muscle DTI. Lastly, our experimental setup did not allow for controlled and standardized force production during muscle contraction. Future studies will investigate the effect of maximal voluntary contraction (MVC) on the diffusion results.

CONCLUSIONS

In conclusion, this study demonstrates that OGSE diffusion encoding allows for quantitative DTI imaging of actively contracting human musculature. The use of cosine trapezoids diffusion encoding waveforms led to reduced signal voids in diffusion-weighted images, which could be used to calculate DTI parameters. Additionally, different diffusion behaviors were observed for relaxed muscle and musculature at different levels of contraction. Taken together, our results showed that OGSE holds the potential to non-invasively assess microstructural changes occurring in the skeletal muscle during contraction and for the non-invasive assessment of contraction abnormalities in patients with muscle disease. These results could lead to a deeper understanding of muscle contraction abnormalities and the optimization of treatment and intervention strategies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors upon request, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Stanford University Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

VM conceived, designed the study, and acquired the data. VM and KM analyzed and interpreted the data. VM drafted the manuscript. KM, FK, BH, and GG revised it. All authors gave their approval of the final submitted version.

FUNDING

This work was supported by Rubicon NWO Grant (452182304) awarded to VM. VM, FK, BH, and GG received research support from GE Healthcare.

REFERENCES

1. Oudeman J, Nederveen AJ, Strijkers GJ, Maas M, Luijten PR, Froeling M. Techniques and applications of skeletal muscle diffusion tensor imaging: A review. *J Magn Reson Imaging*. (2016) 43:773–88. doi: 10.1002/jmri.25016
2. Hooijmans MT, Monte JRC, Froeling M, Berg-faay S Van Den, Aengevaeren VL, Hemke R, et al. Quantitative MRI reveals microstructural changes in the upper leg muscles after running a marathon. *J Magn Reson Imaging*. (2020) 52:407–17. doi: 10.1002/jmri.27106
3. Berry DB, Regner B, Galinsky V, Ward SR, Frank LR. Relationships between tissue microstructure and the diffusion tensor in simulated skeletal muscle. *Magn Reson Med*. (2018) 80:317–29. doi: 10.1002/mrm.26993
4. Mazzoli V, Oudeman J, Nicolay K, Maas M, Verdonschot N, Sprengers AM, et al. Assessment of passive muscle elongation using Diffusion Tensor MRI: correlation between fiber length and diffusion coefficients. *NMR Biomed*. (2016) 29:1813–24. doi: 10.1002/nbm.3661
5. Schwenzer NF, Steidle G, Martirosian P, Schraml C, Springer F, Claussen CD, et al. Diffusion tensor imaging of the human calf muscle: distinct changes in fractional anisotropy and mean diffusion due to passive muscle shortening and stretching. *NMR Biomed*. (2009) 22:1047–53. doi: 10.1002/nbm.1409
6. Sinha S, Sinha U. Reproducibility analysis of diffusion tensor indices and fiber architecture of human calf muscles in vivo at 1.5 Tesla in neutral and plantarflexed ankle positions at rest. *J Magn Reson Imaging*. (2011) 34:107–19. doi: 10.1002/jmri.22596
7. Schwartz M, Steidle G, Martirosian P, Ramos-Murguialday A, Preißl H, Stemmer A, et al. Spontaneous mechanical and electrical activities of human calf musculature at rest assessed by repetitive single-shot diffusion-weighted MRI and simultaneous surface electromyography. *Magn Reson Med*. (2018) 79:2784–94. doi: 10.1002/mrm.26921
8. Schwartz M, Martirosian P, Steidle G, Erb M, Stemmer A, Yang B, et al. Volumetric assessment of spontaneous mechanical activities by simultaneous multi-slice MRI techniques with correlation to muscle fiber orientation. *NMR Biomed*. (2018) 31:e3959. doi: 10.1002/nbm.3959
9. Steidle G, Schick F. Addressing spontaneous signal voids in repetitive single-shot DWI of musculature: Spatial and temporal patterns in the calves of healthy volunteers and consideration of unintended muscle activities as underlying mechanism. *NMR Biomed*. (2015) 28:801–10. doi: 10.1002/nbm.3311
10. Birkbeck MG, Heskamp L, Schofield IS, Blamire AM, Whittaker RG. Non-invasive imaging of single human motor units. *Clin Neurophysiol*. (2020) 131:1399–406. doi: 10.1016/j.clinph.2020.02.004
11. Kim S, Chi-Fishman G, Barnett AS, Pierpaoli C. Dependence on diffusion time of apparent diffusion tensor of ex vivo calf tongue and heart. *Magn Reson Med*. (2005) 54:1387–96. doi: 10.1002/mrm.20676
12. Fieremans E, Lemberskiy G, Veraart J, Sigmund EE, Gyftopoulos S, Novikov DS. In vivo measurement of membrane permeability and myofiber size in human muscle using time-dependent diffusion tensor imaging and the random permeable barrier model. *NMR Biomed*. (2017) 30:e3612. doi: 10.1002/nbm.3612
13. Sigmund EE, Novikov DS, Sui D, Ukpebor O, Baete S, Babb JS, et al. Time-dependent diffusion in skeletal muscle with the random permeable barrier model (RPBM): application to normal controls and chronic exertional compartment syndrome patients. *NMR Biomed*. (2014) 27:519–28. doi: 10.1002/nbm.3087

14. Porcari P, Hall MG, Clark CA, Greally E, Straub V, Blamire AM. The effects of ageing on mouse muscle microstructure: A comparative study of time-dependent diffusion MRI and histological assessment. *NMR Biomed*. (2018) 31:1–13. doi: 10.1002/nbm.3881
15. Karampinos DC, Banerjee S, King KF, Link TM, Majumdar S. Considerations in high-resolution skeletal muscle diffusion tensor imaging using single-shot echo planar imaging with stimulated-echo preparation and sensitivity encoding. *NMR Biomed*. (2012) 25:766–78. doi: 10.1002/nbm.1791
16. Baron CA, Beaulieu C. Oscillating gradient spin-echo (OGSE) diffusion tensor imaging of the human brain. *Magn Reson Med*. (2014) 72:726–36. doi: 10.1002/mrm.24987
17. Parsons EC, Does MD, Gore JC. Temporal diffusion spectroscopy: theory and implementation in restricted systems using oscillating gradients. *Magn Reson Med*. (2006) 55:75–84. doi: 10.1002/mrm.20732
18. Tétreault P, Harkins KD, Baron CA, Stobbe R, Does MD, Beaulieu C. Diffusion time dependency along the human corpus callosum and exploration of age and sex differences as assessed by oscillating gradient spin-echo diffusion tensor imaging. *Neuroimage*. (2020) 210:116533. doi: 10.1016/j.neuroimage.2020.116533
19. Jiang X, Li H, Xie J, McKinley ET, Zhao P, Gore JC, et al. *In vivo* imaging of cancer cell size and cellularity using temporal diffusion spectroscopy. *Magn Reson Med*. (2017) 78:156–64. doi: 10.1002/mrm.26356
20. Van AT, Holdsworth SJ, Bammer R. *In vivo* investigation of restricted diffusion in the human brain with optimized oscillating diffusion gradient encoding. *Magn Reson Med*. (2014) 71:83–94. doi: 10.1002/mrm.24632
21. Veraart J, Novikov DS, Christiaens D, Ades-aron B, Sijbers J, Fieremans E. Denoising of diffusion MRI using random matrix theory. *Neuroimage*. (2016) 142:394–406. doi: 10.1016/j.neuroimage.2016.08.016
22. Klein S, Staring M, Murphy K, Viergever MA, Pluim J. elastix: a toolbox for intensity-based medical image registration. *IEEE Trans Med Imaging*. (2010) 29:196–205. doi: 10.1109/TMI.2009.2035616
23. Veraart J, Sijbers J, Sunaert S, Leemans A, Jeurissen B. Weighted linear least squares estimation of diffusion MRI parameters: Strengths, limitations, and pitfalls. *Neuroimage*. (2013) 81:335–46. doi: 10.1016/j.neuroimage.2013.05.028
24. Tax CMW, Otte WM, Viergever MA, Dijkhuizen RM, Leemans A. REKINDLE: robust extraction of kurtosis indices with linear estimation. *Magn Reson Med*. (2015) 73:794–808. doi: 10.1002/mrm.25165
25. Froeling M. QMRTools: a mathematica toolbox for quantitative MRI analysis. *J Open Source Softw*. (2019) 4:1204. doi: 10.21105/joss.01204
26. Sigmund EE, Baete SH, Luo T, Patel K, Wang D, Rossi I, et al. MRI assessment of the thigh musculature in dermatomyositis and healthy subjects using diffusion tensor imaging, intravoxel incoherent motion and dynamic DTI. *Eur Radiol*. (2018) 28:5304–15. doi: 10.1007/s00330-018-5458-3
27. Whittaker RG, Porcari P, Braz L, Williams TL, Schofield IS, Blamire AM. Functional magnetic resonance imaging of human motor unit fasciculation in amyotrophic lateral sclerosis. *Ann Neurol*. (2019) 85:455–9. doi: 10.1002/ana.25422
28. Oudeman J, Mazzoli V, Marra MA, Nicolay K, Maas M, Verdonschot N, et al. A novel diffusion-tensor MRI approach for skeletal muscle fascicle length measurements. *Physiol Rep*. (2016) 4:e13012. doi: 10.14814/phy2.13012
29. Filli L, Kenkel D, Wurnig MC, Boss A. Diffusional kurtosis MRI of the lower leg: changes caused by passive muscle elongation and shortening. *NMR Biomed*. (2016) 29:767–75. doi: 10.1002/nbm.3529
30. Schlaffke L, Rehmann R, Froeling M, Kley R, Tegenthoff M, Vorgerd M, et al. Diffusion tensor imaging of the human calf: Variation of inter- and intramuscle-specific diffusion parameters. *J Magn Reson Imaging*. (2017) 46:1137–48. doi: 10.1002/jmri.25650
31. Heskamp L, Birkbeck MG, Schofield IS, Whittaker RG, Blamire AM. Detecting motor unit activity during volitional muscle contraction - Motor Unit MRI. *Proc Intl Soc Mag Reson Med*. (2020) 28:2706.
32. Welsh CL, DiBella EV, Hsu E. Higher-order motion-compensation for *in vivo* cardiac diffusion tensor imaging in rats. *IEEE Trans Med Imaging*. (2015) 34:1843–53. doi: 10.1109/TMI.2015.2411571
33. Aliotta E, Wu HH, Ennis DB. Convex optimized diffusion encoding (CODE) gradient waveforms for minimum echo time and bulk motion–compensated diffusion-weighted MRI. *Magn Reson Med*. (2017) 77:717–29. doi: 10.1002/mrm.26166
34. Peña-Nogales Ó, Zhang Y, Wang X, de Luis-Garcia R, Aja-Fernández S, Holmes JH, et al. Optimized Diffusion-Weighting Gradient Waveform Design (ODGD) formulation for motion compensation and concomitant gradient nulling. *Magn Reson Med*. (2019) 81:989–1003. doi: 10.1002/mrm.27462
35. Moulin K, Aliotta E, Ennis DB. Effect of flow-encoding strength on intravoxel incoherent motion in the liver. *Magn Reson Med*. (2019) 81:1521–33. doi: 10.1002/mrm.27490
36. Monte JR, Hooijmans MT, Froeling M, Oudeman J, Tol JL, Maas M, et al. The repeatability of bilateral diffusion tensor imaging (DTI) in the upper leg muscles of healthy adults. *Eur Radiol*. (2020) 30:1709–18. doi: 10.1007/s00330-019-06403-5
37. Wu D, Zhang J. The effect of microcirculatory flow on oscillating gradient diffusion MRI and Diffusion Encoding with Dual-Frequency Orthogonal Gradients (DEFOG). *Magn Reson Med*. (2017) 77:1583–92. doi: 10.1002/mrm.26242

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fneur.2021.632096

Nerve Ultrasound Comparison Between Transthyretin Familial Amyloid Polyneuropathy and Chronic Inflammatory Demyelinating Polyneuropathy

Kang Du[1†], Ke Xu[1†], Si Cheng[1,2†], He Lv[1], Wei Zhang[1], Zhaoxia Wang[1], Yun Yuan[1] and Lingchao Meng[1]*

[1] Department of Neurology, Peking University First Hospital, Beijing, China, [2] Beijing Tiantan Hospital, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Jordi Diaz-Manera,
University of Newcastle,
United Kingdom

Reviewed by:
Paola Sandroni,
Mayo Clinic, United States
Alberto Stefano Tagliafico,
University of Genoa, Italy

***Correspondence:**
Lingchao Meng
lcmeng@bjmu.edu.cn

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: 22 November 2020
Accepted: 08 February 2021
Published: 26 February 2021

Citation:
Du K, Xu K, Cheng S, Lv H, Zhang W, Wang Z, Yuan Y and Meng L (2021) Nerve Ultrasound Comparison Between Transthyretin Familial Amyloid Polyneuropathy and Chronic Inflammatory Demyelinating Polyneuropathy. Front. Neurol. 12:632096. doi: 10.3389/fneur.2021.632096

Backgrounds: Transthyretin familial amyloid polyneuropathy (TTR-FAP) is frequently misdiagnosed as chronic inflammatory demyelinating polyneuropathy (CIDP) because of similar phenotypes in the two diseases. This study was intended to identify the role of nerve ultrasonography in evaluating TTR-FAP and CIDP.

Methods: Eighteen patients with TTR-FAP, 13 patients with CIDP, and 14 healthy controls (HC) were enrolled in this study. Consecutive ultrasonography scanning was performed in six pairs of nerves of bilateral limbs with 30 sites. The cross-sectional areas (CSAs) and CSA variability data of different groups were calculated and compared.

Results: Both TTR-FAP and CIDP showed larger CSAs at most sites of both upper and lower limbs than in HC groups. CIDP patients had larger CSAs than TTR-FAP patients at 8/15 of these sites, especially at U1-3, Sci2 sites ($p < 0.01$). However, the CSAs at above sites were not a credible index to differentiate TTR-FAP from CIDP with a low area under the curve (<0.8). The CSA variability of median nerves was significantly higher in CIDP than in TTR-FAP and HC groups, with high sensitivity (0.692) and specificity (0.833) to differentiate CIDP from TTR-FAP. The CSA variability of ulnar nerves was not significantly different between the three groups. For the TTR-FAP group, mean CSAs at each site were not correlated with different Coutinho stages, modified polyneuropathy disability, course of sensory motor peripheral neuropathy, Neuropathy Impairment Score, or Norfolk Quality of life-diabetic neuropathy score. The mean compound muscle action potential of ulnar nerves was negatively correlated with the mean CSAs of ulnar nerves.

Interpretation: TTR-FAP patients had milder nerve enlargement with less variability in CSAs of median nerves than those with CIDP, suggesting that nerve ultrasound can be a potential useful auxiliary tool to help differentiate the two neuropathies.

Keywords: TTR-FAP, CIDP, transthyretin, polyneuropathy, ultrasonography

INTRODUCTION

Transthyretin familial amyloid neuropathy (TTR-FAP) is a multiple systemic disorder caused by *TTR* gene mutation and characterized by extracellular deposition of transthyretin-derived amyloid fibrils in peripheral and autonomic nerves and other organs. The typical phenotype of TTR-FAP is severe progressive sensory and motor neuropathy with autonomic neuropathy among adults, and most of them with cardiomyopathy. The pathology of TTR-FAP is characterized by TTR deposition with diffuse loss of nerve fibers. However, phenotypic variability and non-disease-specific symptoms or unknown family history often delay diagnosis and lead to misdiagnosis (1), including chronic inflammatory demyelinating polyneuropathy (CIDP). Some sporadic cases present with the demyelinating process in nerve conduction studies (NCSs) (2, 3), which fulfills both clinical and electrophysiologic criteria for CIDP during initial evaluation (4). Since early differentiation of TTR-FAP from CIDP is important for the treatment of either disease, several electrophysiological studies were performed for differential diagnosis. Quantitative sudomotor test was used to distinguish CIDP from TTR-FAP with good sensitivity and specificity (5).

Nerve ultrasound is a painless tool for quick evaluation of peripheral nerve morphology. Several nerve ultrasound studies showed nerve enlargement in TTR-FAP (6–8). The cross-sectional areas (CSAs) of peripheral nerves in cases of TTR-FAP are significantly larger than those of controls, most are in the proximal nerve segments (7). Nerve ultrasound patterns can facilitate the evaluation of asymptomatic carriers, presenting as larger nerve CSAs at proximal nerve sites (8). Nerve ultrasound can also serve as a useful complementary diagnostic tool for the identification of treatment-responsive inflammatory neuropathies (9–11). Sonographic nerve enlargement was present in all patients and was most prominent in proximal segments of the median nerve and brachial plexus (9, 10), including the fascicle CSAs in CIDP (12). The nerve ultrasound finding of CIDP is different from that of demyelinating diabetic sensorimotor polyneuropathy (9). However, there has been no study so far on the nerve ultrasound comparison between CIDP and TTR-FAP. In this study, more unabridged nerve sites including both upper and lower limbs were measured and the CSA variability of CIDP and TTR-FAP patients was compared.

MATERIALS AND METHODS

Subjects

Between June 2015 and September 2020, 18 patients (16 males and 2 females) with TTR-FAP, 13 patients (3 males and 10 females) with CIDP, and 14 healthy controls (8 males and 6 females) were recruited in Peking University First Hospital. All TTR-FAP patients were diagnosed according to the diagnostic criteria (1). For the diagnosis of definite CIDP, the diagnostic criteria proposed by the Joint Task Force of the European Federation of Neurological Societies and the Peripheral Nerve Society (EFNS/PNS) were used (4). The exclusion criteria of healthy controls were: [1] skin numbness or paresthesia; [2] muscle atrophy or weakness; [3] other disorders of the peripheral nervous system; and [4] chronic diseases of other organs (e.g., heart, brain, eye, and kidney).

The mean age of TTR-FAP patients, CIDP patients, and healthy controls was 45.8 years (range 26–64 years), 40.7 years (range 15–69 years), and 40.3 years (range 26–65 years), respectively. There was no statistically significant difference in age between the three groups ($p = 0.587$).

Clinical Neurologic Evaluation of TTR-FAP and CIDP Patients

All TTR-FAP subjects diagnosed with mutations in the *TTR* gene were inquired about their disease history and had a focused neurological examination of measurement scales performed, including Neuropathy Impairment Score (NIS), Norfolk Quality of life-diabetic neuropathy score (Norfolk QOL-DN), and modified polyneuropathy disability (m-PND). Disease severity was estimated by Coutinho staging of TTR-FAP. Nerve conduction studies (NCSs) were performed in all TTR-FAP patients according to the standard protocol using surface stimulation and recording. The motor nerve conduction velocity (MCV) and distal compound muscle action potential (CMAP) of the bilateral median, ulnar nerves of 11 patients were included in this study. All CIDP subjects were asked about their detailed disease history, and underwent neurological examination and NCSs. Sural nerve biopsy was conducted for 17/18 of TTR-FAP patients and most of the CIDP patients (9/13). Congo red staining and TTR immunohistochemical staining were performed in 17/18 and 13/18 of the TTR-FAP patients, respectively.

Ultrasonographic Studies

All subjects underwent peripheral nerve ultrasound using the Philips Imaging System (iU Elite, Bothell, WA, USA) that measured and recorded the bilateral median, ulnar, sciatic, tibial, common peroneal, and sural nerves. To be more specific, the 17MHz high-frequency linear array probe was used for the superficial nerves, including the median nerves, ulnar nerves, common peroneal nerves, and sural nerves, and the 9 MHz linear array probe was used for the deeper nerves, including the sciatic nerves and tibial nerves.

The CSAs at the predetermined sites of each nerve were measured by tracing just inside the hyperechoic rim of the nerve. Thirty predetermined sites were measured of all the nerves (13), including [1] 10 sites that were measured in left and right median nerves (LM & RM): LM1/RM1= wrist (entrance of the carpal tunnel at the pisiform bone level); LM2/RM2= distal forearm (the nerve reached the deep flexor digitorum and started to traverse between the deep flexor digitorum and the flexor pollicis longus); LM3/RM3= proximal forearm (the clearest point before the nerve entered pronator teres); LM4/RM4= elbow (elbow socket); LM5/RM5= upper arm (from cubital fossa to the middle of armpit). [2] 10 sites of left and right ulnar nerves (LU & RU): LU1/RU1= wrist (Guyon tube: between nerve deviation and the pisiform bone and ulnar artery); LU2/RU2= distal forearm (before the ulnar nerve branches off); LU3/RU3= proximal forearm (2/3 between the wrist and elbow); LU4/RU4= elbow (at the medial epicondyle of humerus); LU5/RU5= upper arm

TABLE 1 | Comparison of CSAs at different nerve sites of upper and lower limbs in TTR-FAP, CIDP and healthy controls (mm^2).

Sites	Mean CSAs (mm^2) of HC	Mean CSAs (mm^2) of TTR-FAP	Mean CSAs (mm^2) of CIDP	*P* value (HC vs. TTR-FAP)	*P* value (TTR-FAP vs. CIDP)	*P* value (HC vs. CIDP)
M1	8.26 (1.65)	11.46 (3.19)	11.1 (3.83)	**0.000****	0.628	**0.001****
M2	7.31 (1.36)	9.04 (2.11)	15.11 (10.07)	**0.001****	**0.029***	**0.000****
M3	7.21 (1.69)	9.57 (2.63)	14.26 (10.68)	**0.000****	0.180	**0.000****
M4	9.12 (1.97)	10.98 (3.07)	15.29 (8.32)	0.065	**0.015***	**0.000****
M5	9.11 (1.92)	13.19 (3.82)	24.84 (18.45)	**0.000****	0.058	**0.000****
U1	5.04 (1.74)	4.84 (1.56)	6.37 (2.67)	0.456	**0.004****	**0.049***
U2	5.95 (1.26)	5.74 (1.72)	9.86 (5.02)	0.267	**0.000****	**0.005****
U3	5.83 (1.39)	6.56 (1.97)	10.78 (6.92)	0.112	**0.001****	**0.000****
U4	8.21 (1.97)	9.56 (4.07)	12.2 (8.52)	0.583	0.284	0.073
U5	6.52 (2.01)	8.77 (3.34)	15.76 (15.25)	**0.003****	**0.049***	**0.000****
Sci1	54.15 (15.7)	76.68 (19.59)	82.36 (39.19)	**0.000****	0.936	**0.002****
Sci2	54.67 (14.49)	70.09 (20.31)	118.92 (70.48)	**0.001****	**0.005****	**0.000****
Pc	11.49 (3.79)	13.83 (3.65)	19.63 (11.93)	**0.024***	**0.022***	**0.001****
Tib	32.36 (7.58)	49.09 (12.63)	47.57 (17.08)	**0.000****	0.476	**0.000****
Sural	5.16 (1.29)	4.62 (1.59)	5.31 (1.42)	0.075	0.064	0.590

*Mean (SD). *Significance at 0.05 level.*
***Significance at 0.01 level.*
All significant p values are printed in bold, and discriminative sites of TTR-FAP and CIDP groups in italics plus bold.
CSAs, cross-sectional areas; CIDP, chronic inflammatory demyelinating polyneuropathy; TTR-FAP, transthyretin familial amyloid polyneuropathy; HC, healthy control.
M, median nerve; U, ulnar nerve; Sci, sciatic nerve; Tib, tibial nerve; Pc, common peroneal nerve; Sural, sural nerve.

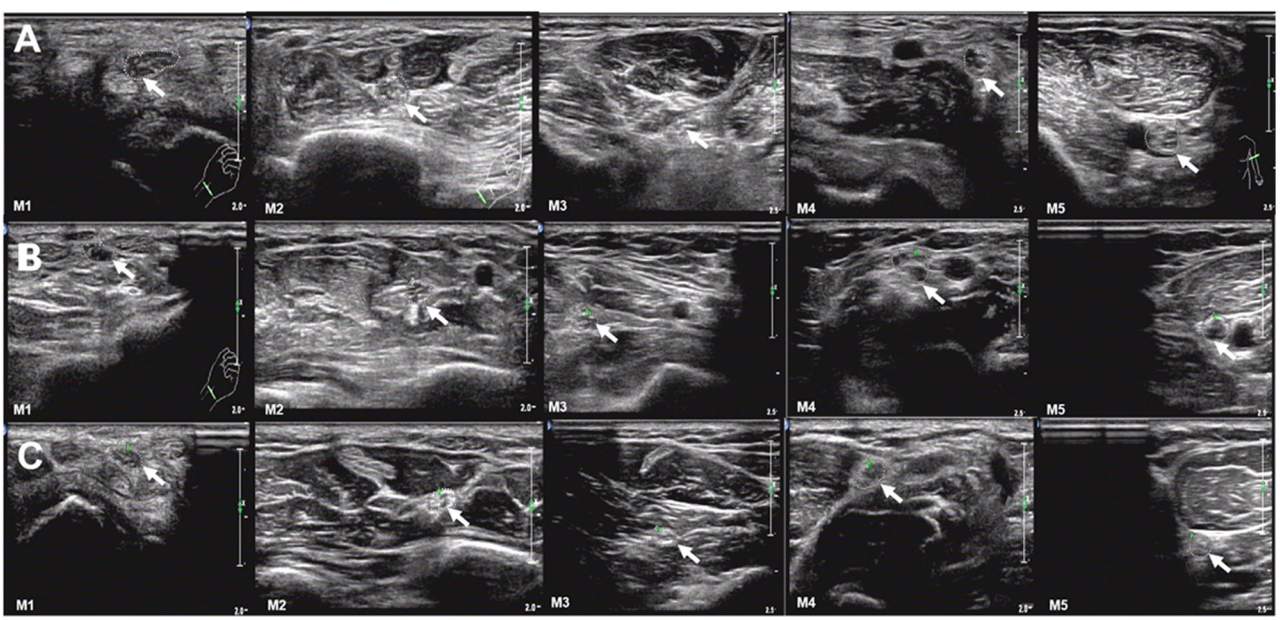


FIGURE 1 | Examples of ultrasound cross-sections showing measurements of cross-sectional area of median nerves in M1-5 between the three groups. **(A)** Homogeneous enlargement was observed at sites M1-5 of a TTR-FAP patient. **(B)** Segmental enlargement was observed at sites M1-5 of median nerve in a CIDP patient. **(C)** The normal CSAs at sites M1-5 of a healthy control.

(from cubital fossa to the middle of armpit). [3] 4 sites in left and right sciatic nerves (LSci & RSci): LSci1/RSci1= middle thigh; LSci2/RSci2= 1/3 of mid-lower part of the thigh (before sciatic nerves were divided into common peroneal nerves and tibial nerves). [4] 2 sites in left and right tibial nerves (LTib & RTib): LTib/ RTib = popliteal fossa (just after the tibial nerves were branched off by sciatic nerves). [5] 2 sites in left and right common peroneal nerves (LPc & RPc): LPc/ RPc=

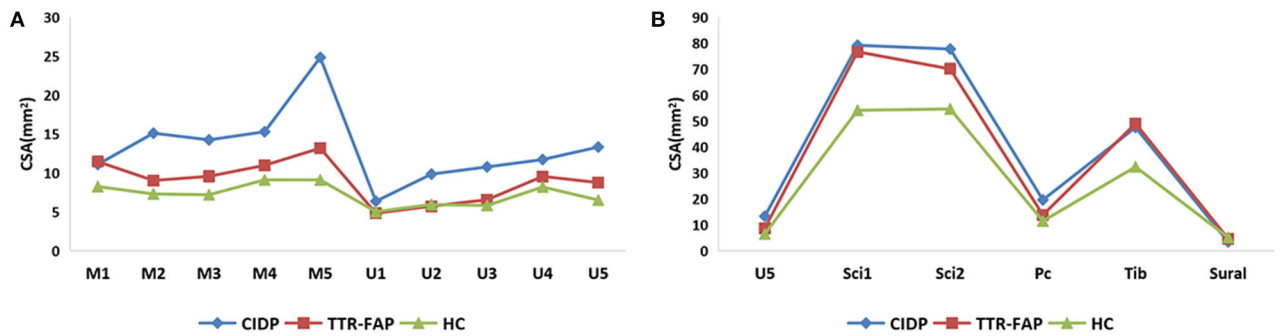


FIGURE 2 | A general view of mean CSAs at sites of upper limbs **(A)** and lower limbs **(B)** between the three groups. **(A)** Clearly presented the CSAs in upper limbs, and revealed most CSAs of TTR-FAP groups at bilateral median nerves were higher than those of HC groups but lower than those of CIDP groups. **(B)** Showed most CSAs of TTR-FAP groups were between those of CIDP and HC groups in lower limbs.

capitulum fibulae. [6] 2 sites of left and right sural nerves (LSural & RSural): Lsual/Rsual = lower 1/4 of the lower leg near lateral malleolus). Left sural and right sural nerves in the TTR-FAP group lacked 11 and 4 CSA data due to sural nerves biopsy, respectively. Left sural and right sural nerves in the CIDP group lacked 5 and 4 CSA data due to sural nerves biopsy, respectively.

The measured parameters were nerve CSAs and CSA variability. The CSAs were measured at these sites of each limb. The CSA variability was defined as "maximum CSA/minimum CSA".

Statistical Analysis

IBM SPSS Statistics, version 26 was used for statistical analysis. The CSAs of healthy controls showed a normal distribution, while those of CIDP and TTR-FAP showed an abnormal distribution (as evaluated by single sample K-S test). Thus, Mann-Whitney *U* test was used for evaluating differences in CSAs between TTR-FAP, CIDP and healthy control groups, as well as CSA variability between TTR-FAP and CIDP groups. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the applicability of CSA variability measurements to differentiation of TTR-FAP from CIDP. The area under the curve (AUC) was calculated. The value of Youden index at its maximum was taken as the cut point for the diagnosis of TTR-FAP, and the sensitivity and specificity were calculated. Two-sided p values were calculated for all analyses; $p < 0.05$ was considered significant. Spearman analysis was used to test the correlation between CSAs and measuring scales, electrophysiological data.

Data Availability

Anonymized data will be shared by request from any qualified investigator.

RESULTS

Clinical Data of TTR-FAP and CIDP

Of the 18 TTR-FAP patients, 9 initially developed limb paresthesia, followed by other onset symptoms such as alternating diarrhea and constipation (ADC) in three patients, sexual dysfunction in three patients, blurred vision in two patients and constipation in one patient. All these patients presented with sensorimotor peripheral neuropathy and autonomic neuropathy, 12 suffered from asymptomatic cardiac hypertrophy, and 5 developed vitreous opacity. *TTR* gene screening was performed, with Val30Met mutation in three patients, Ala97Ser, Glu42Gly, Gly47Arg, and Lys35Asn mutation in two patients, respectively, Ala36Pro, Phe33Leu, Phe33Val, Gly83Arg, Ser77Phe, Ser77Tyr, and Val28Ser mutation in one patient, respectively. NCSs examination was also performed in all these patients: 12 presented with axonal impairment, and 6 with a mixed neuropathy. In clinical staging, 10 of these patients were divided into Coutinho stage I, and the remaining into Coutinho stage II or III.

For CIDP patients, proximal and/or distal limbs weakness were manifested, with or without paresthesia. The mean course of disease was 3.4 ± 2.2 years. Laboratory examination of cerebrospinal fluid was conducted in 8/13 of these patients with cytoalbuminologic dissociation. All the patients undergoing NCSs examination accorded with the presentation of demyelination. All patients received immunotherapy that turned out to be partly or completely effective.

Pathologically, positive Congo red staining in sural nerve biopsy was seen in 10/17 of TTR-FAP patients, positive TTR immunohistochemistry in 6/13 of these patients. All patients with nerve biopsy pathologically presented with axonal neuropathy with moderate to severe loss of both large and small myelinated nerve fibers as well as unmyelinated nerve fibers. Of the 9 CIDP patients who had sural nerve biopsy performed, variation in the density of myelinated fibers among fascicles was observed in 5 patients, infiltration of macrophage in 5 patients, and thin myelin sheath or onion-bulb formation in 6 patients. All these patients had mild to moderate loss of myelinated fibers, especially large-diameter ones.

Ultrasonographic Findings

Comparison of CSAs Between TTR-FAP, CIDP, and HC

The mean CSAs at 15 different sites of all nerves in each group were shown in **Table 1**. The mean CSAs values in the TTR-FAP

TABLE 2 | Comparison of CSA variability of median/ulnar nerve between disease groups and control group.

	TTR-FAP	CIDP	HC	*P* value		
				TTR-FAP v.s. CIDP	TTR-FAP v.s. HC	CIDP v.s. HC
M-CSA-V	1.58 (0.32)	3.06 (1.88)	1.59 (0.44)	**0.000****	0.074	**0.000****
U-CSA-V	2.14 (0.47)	3.22 (2.50)	1.95 (0.54)	0.608	0.823	0.057

Mean (SD).
***Significant difference at 0.01 level.*
M/U-CSA-V: CSA variability of median/ulnar nerve, defined as "maximum CSA/ minimum CSA". All significant p values are printed in bold.

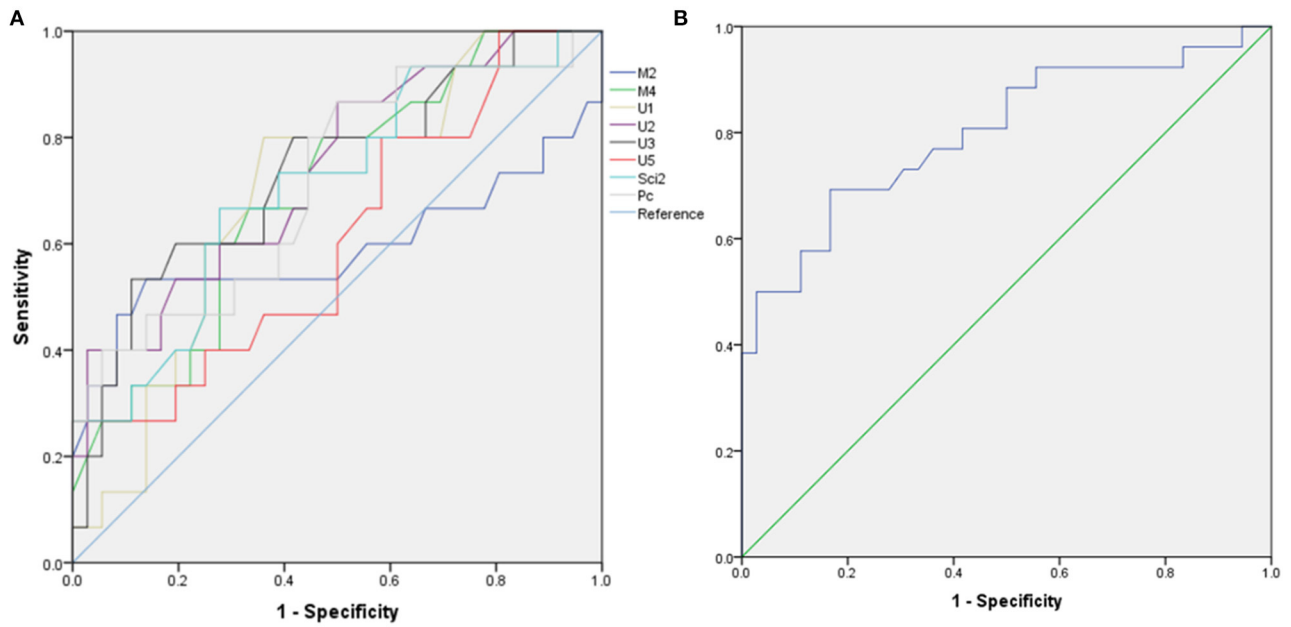


FIGURE 3 | Receiver operating characteristic (ROC) curve for differentiating transthyretin familial amyloid polyneuropathy (TTR-FAP) from chronic inflammatory demyelinating polyneuropathy (CIDP). **(A)** Mean CSAs at M2, M4, U1-3, U5, Sci2, Pc sites between the two groups. **(B)** CSA-variability of median nerves between the two groups.

group were statistically higher than those of the HC group at most sites of median, sciatic, tibial nerves, especially in median nerves and sciatic nerves (all $p < 0.05$), including M1-M3, M5, U5, Sci1, Sci2, Pc, Tib sites. The CSAs at most sites of ulnar nerves were not higher than those of the HC groups, except U5 site, which was the proximal site of ulnar nerves. The CSAs of proximal sites of median nerves (M5) and sciatic nerves (Sci1) in the TTR-FAP group were also significantly higher than in the HC group (**Figures 1**, **2**).

The mean CSAs values at 8 sites of the TTR-FAP group were lower than in CIDP with significant difference, including M2, M4, U1-3, U5, Sci2, Pc sites. For CIDP groups, the CSAs at all sites were higher in HC groups intuitively, but were not significantly different at two sites (U4 and Sural) (**Figures 1**, **2**).

Comparison of CSA Variability Between TTR-FAP and CIDP

Furthermore, the CSA variability of median nerves and ulnar nerves between the three groups was calculated. It was found that CSA variability of median nerves in CIDP groups was significantly higher than in TTR-FAP and HC groups, but there was no significant difference between TTR-FAP and HC groups. For ulnar nerves, the CSA variability between the three groups was not significantly different (**Table 2**).

The ROC of CSA and CSA Variability for Differentiating Between TTR-FAP and CIDP

Based on the results observed in **Tables 1**, **2**, we went to performed the ROC curve of CSAs and CSA variability for differentiating between TTR-FAP and CIDP. **Figure 3** showed the ROC curve analyses of the mean CSAs of the discriminative sites in each nerve and CSA variability in median nerves. AUC and cutoff values were shown in **Table 3**. There was no significant difference between the two groups in CSAs at M2 and U5 sites ($p > 0.05$), except M4, U1-3, U5, Sci2, Pc sites. However, the AUC above was not high (all AUC<0.8). For CSA variability in median nerves, the AUC was 0.8 with high sensitivity (0.692) and specificity (0.833).

TABLE 3 | The AUC, suggested cutoff values, sensitivity, and specificity of CSA and CSA variability in differentiating between TTR-FAP and CIDP.

	AUC	Cutoff values	Sensitivity	Specificity	*p* value
M-CSA-V	0.8	1.77	0.692	0.833	**0.000****
M2	0.59	11.95	0.533	0.861	0.316
M4	0.699	10.75	0.667	0.667	**0.026***
U1	0.695	4.45	0.8	0.639	**0.029***
U2	0.736	9.2	0.4	0.972	**0.008****
U3	0.734	8.2	0.533	0.889	**0.009****
U5	0.61	19.45	0.267	1	0.219
Sci2	0.707	78.8	0.667	0.722	**0.021***
Pc	0.715	13.15	0.867	0.5	**0.016***

***Significant difference at 0.01 level.*
**Significant difference at 0.05 level.*
AUC, area under the curve; M-CSA-V, CSA-Variability of median nerve. All significant p values are printed in bold.

TABLE 4 | Comparison of CSAs in different Coutinho stages of each measurement site in TTR-FAP group.

Sites	Mean CSA of Stage I (mm^2)	Mean CSA of Stage II/III (mm^2)	*P* value
M1	11.54 (3.23)	11.36 (3.25)	0.69
M2	9 (2.16)	9.09 (2.1)	0.949
M3	9.52 (3.02)	9.64 (2.14)	0.774
M4	11.19 (3.48)	10.73 (2.54)	0.987
M5	13.21 (3.7)	13.16 (4.09)	0.691
U1	4.38 (0.84)	5.41 (2.04)	0.299
U2	5.38 (1.38)	6.18 (2.02)	0.134
U3	6.23 (1.9)	6.98 (2.04)	0.082
U4	9.23 (3.81)	9.98 (4.46)	0.474
U5	8.34 (3.56)	9.3 (3.06)	0.119
Sci1	78.51 (22.7)	74.4 (15.25)	0.924
Sci2	70.93 (24.28)	69.05 (14.6)	0.524
Pc	14.35 (3.53)	13.18 (3.8)	0.339
Tib	48.24 (12.8)	50.15 (12.75)	0.373
Sural	4.27 (1.03)	5.1 (2.1)	0.455

Mean (SD).

Correlation of CSAs With Electrophysiological Data and Clinical Measurement Scales of TTR-FAP

All the measured sites between Coutinho stage I ($n = 10$) and Coutinho stage II/III ($n = 8$) of TTR-FAP patients were compared. However, the CSAs of these two groups at each site were not significantly different (all $p > 0.05$) (**Table 4**).

In addition, the correlation analysis was conducted of mean CSAs in median and ulnar nerves and of Neuropathy Impairment Score (NIS) that included one item for reflection of muscle weakness, modified polyneuropathy disability (m-PND), Norfolk Quality of life-diabetic neuropathy score (Norfolk QOL-DN), the course of sensory motor peripheral neuropathy (SMPN) and electrophysiological data in TTR-FAP patients. However, no correlation was observation between all these indexes (**Table 5**), except the mean CMAP of ulnar nerves, which was negatively correlated with mean CSAs with statistically significant difference ($r = -0.491$, $p = 0.008$) (**Figure 4**).

DISCUSSION

The clinical symptoms and electrophysiology might be similar in TTR-FAP and CIDP patients. In TTR-FAP, destruction of myelin due to amyloid deposition might be related to nerve conduction abnormalities mimicking CIDP (5). Initial electrodiagnostic conclusions of CIDP were confirmed in only 45% of misdiagnosed studies (14).

Our study conformed with the findings of previous studies (6, 7, 15, 16) that thickness of peripheral nerves existed in TTR-FAP patients. The mean CSAs of TTR-FAP patients were higher than those of healthy controls at most sites. Enlargement of peripheral nerves has been reported in previous studies (6, 7). The CSAs at proximal sites of measurable nerves (median nerves, ulnar nerves and sciatic nerves) were significantly higher than those of healthy controls, compared with distal sites in a same nerve. All this was compatible with the findings of previous studies for nerve ultrasound and magnetic resonance neurography in TTR-FAP patients (7, 8, 15). Moreover, prominent enlargement of peripheral nerves at proximal sites was not common in most axonal neuropathies (17, 18), which could help distinguish TTR-FAP from other axonal neuropathies. For note, distal enlargement of median nerves was also observed in our study, which might associate with carpal tunnel syndrome in these TTR-FAP patients (8).

To our knowledge, there were no studies on how nerve ultrasound was used for comparing CIDP with TTR-FAP. A recent case report on TTR-FAP said that the CSAs were not different between CIDP patients and those with TTR-FAP. Instead, nerve ultrasound features of TTR-FAP could increase the incidence of misdiagnosis of CIDP. However, only one TTR-FAP patient was involved in this case (16). We found different results of CSAs between the two groups, and revealed that CSAs of enlarged peripheral nerves of TTR-FAP patients were lower than those of CIDP patients at 8/15 of sites with significant difference, especially at sites of U1-3, Sci2 (**Table 1**). A second point that had been neglected by other studies was the fluctuation of different CSAs in a same nerve (i.e., median and ulnar nerves) among patients with CIDP (19), which might be an auxiliary index for differentiating

TABLE 5 | Correlation analysis between CSAs at different sites and clinical as well as electrophysiological data of TTR-FAP patients.

	Mean CSAs of median nerves	Mean CSAs of ulnar nerves	Mean CSAs of sciatic nerves	Mean CSAs of common peroneal nerves	Mean CSAs of tibial nerves	Mean CSAs of sural nerves	Mean CSAs at M-5 sites	Mean CSAs at Sci-1 sites
MCV of median nerves	$r = 0.133$, $p = 0.469$	NA	NA	NA	NA	NA	$r = 0.097$, $p = 0.596$	NA
Mean CMAP of median nerves	$r = 0.167$, $p = 0.36$	NA	NA	NA	NA	NA	$r = 0.177$, $p = 0.331$	NA
MCV of ulnar nerves	NA	$r = -0.095$, $p = 0.645$	NA	NA	NA	NA	NA	NA
Mean CMAP of ulnar nerves	NA	$r = -0.491$, **$p = 0.008$****	NA	NA	NA	NA	NA	NA
MCV of common peroneal nerves	NA	NA	NA	$r = -0.255$, $p = 0.326$	NA	NA	NA	NA
Mean CMAP of common peroneal nerves	NA	NA	NA	$r = -0.34$, $p = 0.096$	NA	NA	NA	NA
MCV of tibial nerves	NA	NA	NA	NA	$r = -0.457$, $p = 0.056$	NA	NA	NA
Mean CMAP of tibial nerves	NA	NA	NA	NA	$r = -0.126$, $p = 0.565$	NA	NA	NA
Course of SMPN	$r = -0.015$, $p = 0.954$	$r = 0.351$, $p = 0.153$	$r = 0.052$, $p = 0.838$	$r = 0.012$, $p = 0.961$	$r = 0.259$, $p = 0.299$	$r = 0.035$, $p = 0.895$	$r = -0.026$, $p = 0.919$	$r = 0.109$, $p = 0.668$
NIS	$r = 0.026$, $p = 0.919$	$r = 0.408$, $p = 0.093$	$r = 0.071$, $p = 0.779$	$r = -0.058$, $p = 0.82$	$r = 0.337$, $p = 0.172$	$r = -0.079$, $p = 0.764$	$r = -0.043$, $p = 0.864$	$r = 0.072$, $p = 0.776$
m-PND	$r = 0.092$, $p = 0.717$	$r = 0.463$, $p = 0.053$	$r = 0.115$, $p = 0.651$	$r = 0.005$, $p = 0.984$	$r = 0.148$, $p = 0.559$	$r = 0.037$, $p = 0.888$	$r = 0.057$, $p = 0.823$	$r = 0.125$, $p = 0.622$
Norfolk QOL-DN	$r = -0.045$, $p = 0.861$	$r = 0.396$, $p = 0.104$	$r = -0.012$, $p = 0.962$	$r = -0.17$, $p = 0.501$	$r = 0.01$, $p = 0.969$	$r = 0.159$, $p = 0.542$	$r = -0.098$, $p = 0.699$	$r = -0.01$, $p = 0.969$
NIS-muscle weakness	$r = 0.02$, $p = 0.938$	$r = 0.324$, $p = 0.19$	$r = 0.009$, $p = 0.971$	$r = -0.091$, $p = 0.719$	$r = 0.302$, $p = 0.223$	$r = -0.077$, $p = 0.769$	$r = -0.078$, $p = 0.759$	$r = 0.004$, $p = 0.987$

Spearman analysis was used to test the correlation between CSAs and electrophysiology, as well as measuring scales.

***Significant correlation at 0.01 level.*

All significant p values are printed in bold.

NA, not available; M-5, site-5 of bilateral median nerves; Sci, site-1 of bilateral sciatic nerves; SMPN, sensory motor peripheral neuropathy; NIS, Neuropathy Impairment Score; NIS-muscle weakness, Neuropathy Impairement Score only including the items of muscle weakness; m-PND, modified polyneuropathy disability; Norfolk QOL-DN, Norfolk Quality of life-diabetic neuropathy score.

CIDP and TTR-FAP. The CSAs of median nerves in TTR-FAP patients were not all significantly higher than in CIDP patients, but CSA variability of the median nerves might help to differentiate CIDP from TTR-FAP due to its relatively high sensitivity and specificity. It was also speculated that the CSAs of ulnar nerves might be a potentially useful indicator for differentiating CIDP from TTR-FAP, unlike the CSAs variability of ulnar nerves.

The nerve ultrasound results may be based on pathological changes. The TTR-FAP was an axonal neuropathy and the loss of nerve fibers was diffuse and regular (20). The possible pathophysiological mechanisms have been clarified as amyloid deposits-vulnerable to compression-compression sites edema, fibrosis, thickened endoneurium, perineurium and the small vessel walls, as well as nerve fiber degeneration (7, 21). CIDP was a demyelination neuropathy characterized by infiltration of macrophage and variation in myelinated fiber density among fascicles due to focal myelinated fiber loss or onion-bulb formation (22, 23).

Previous studies suggested that disease duration, stage of TTR-FAP, or PND stage were not correlated with CSAs of median nerves (7), which was why we evaluated ulnar nerves more comprehensively. Similarly, we confirmed that disease severity, including m-PND and Coutinho staging, was not associated with CSAs in TTR-FAP patients in our study. The correlation between NIS, Norfolk QOL-DN and mean CSAs of each nerve was not observed.

Interestingly, the mean CMAP of ulnar nerves was negatively correlated with the mean CSAs in our study, suggesting that the CSAs of ulnar nerves might be used to monitor the disease severity, but further studies are needed. The negative correlation between CMAP and CSAs was observed in median nerves in previous studies (7), but not in our current study.

This study had several limitations. The sample size of TTR-FAP patients with electrophysiological data and CIDP patients was not big enough, so more subjects registered will be needed in the future. CIDP itself is a heterogeneous disease, so the CSA of each site may be affected by different

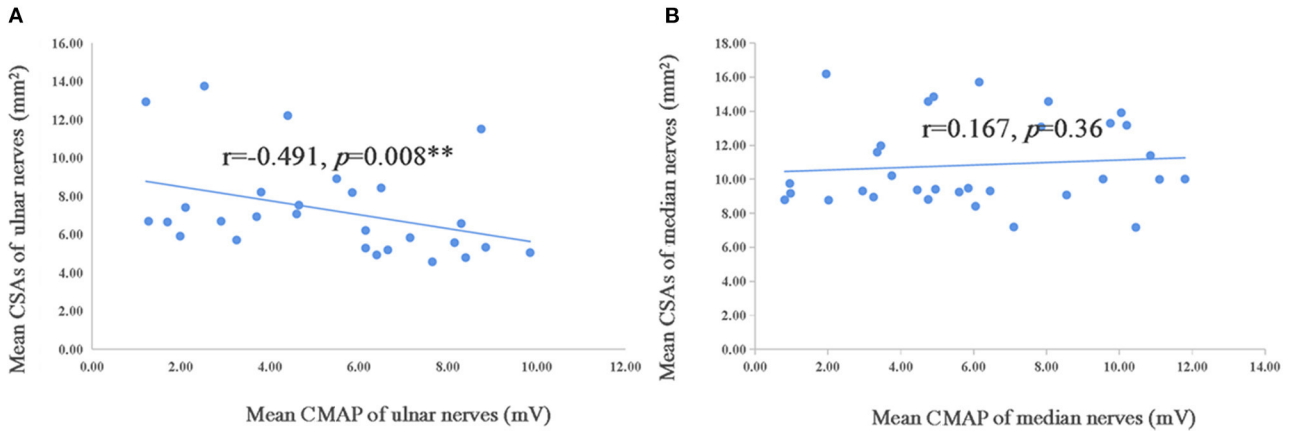


FIGURE 4 | Correlation of mean CSAs and CMAP of ulnar and median nerves in TTR-FAP group. The Spearman correlation coefficient was calculated. **Significance at 0.01 level. CMAP, compound motor action potential.

disease subtypes and activities. Compared with pathological examination and genetic testing, which are the golden standard to differentiate TTR-FAP and CIDP, nerve ultrasound can only be considered as an auxiliary tool, with the non-invasive and convenient advantages.

In conclusion, our study showed TTR-FAP patients had milder nerve enlargement with less variability in CSAs of median nerves than their CIDP counterparts, suggesting that nerve ultrasound is a potential useful auxiliary tool in differentiating the two neuropathies.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the clinical research ethics committee of Peking University First Hospital. The patients/participants provided written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KD: acquisition of data, completion of statistical analysis, and drafting of the initial manuscript and writing of the final manuscript. KX and SC: ultrasonography, study concept and design, and critical revision of the manuscript. HL, WZ, and YY: study concept and design, and critical revision of the manuscript. LM: data review, interpretation of results, and revision of the initial draft. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Beijing Municipal Natural Science Foundation (7194323) and Peking University Medicine Fund of Fostering Young Scholar's Scientific & Technological Innovation (BMU2020PY011).

REFERENCES

1. Adams D, Ando Y, Beirão JM, Coelho T, Gertz MA, Gillmore JD, et al. Expert consensus recommendations to improve diagnosis of ATTR amyloidosis with polyneuropathy. *J Neurol*. (2020). doi: 10.1007/s00415-019-09688-0
2. Lozeron P, Mariani L, Dodet P, Beaudonnet G, Théaudin M, Adam C, et al. Transthyretin amyloid polyneuropathies mimicking a demyelinating polyneuropathy. *Neurology*. (2018) 91(2):e143-e152. doi: 10.1212/WNL.0000000000005777
3. Ohashi N, Kodaira M, Morita H, Sekijima Y. Electrophysiological demyelinating features in hereditary ATTR amyloidosis. *Amyloid*. (2019) 26:15-23. doi: 10.1080/13506129.2018.1564903
4. Van den Bergh PYK, Hadden RDM, Bouche P, Cornblath DR, Hahn A, Illa I, et al. European Federation of Neurological Societies/Peripheral Nerve Society Guideline on management of chronic inflammatory demyelinating polyradiculoneuropathy: Report of a joint task force of the European Federation of Neurological Societies and the Peripher. *Eur J Neurol*. (2010) 17:356–63. doi: 10.1111/j.1468-1331.2009.02930.x
5. Fortanier E, Delmont E, Verschueren A, Attarian S. Quantitative sudomotor test helps differentiate transthyretin familial amyloid polyneuropathy from chronic inflammatory demyelinating polyneuropathy. *Clin Neurophysiol*. (2020) 131:1129–33. doi: 10.1016/j.clinph.2020.01.022
6. Granata G, Luigetti M, Coraci D, Del Grande A, Romano A, Bisogni G, et al. Ultrasound evaluation in transthyretin-related amyloid neuropathy. *Muscle Nerve*. (2014) 50:372–6. doi: 10.1002/mus.24168
7. Podnar S, Sarafov S, Tournev I, Omejec G, Zidar J. Peripheral nerve ultrasonography in patients with transthyretin amyloidosis. *Clin Neurophysiol*. (2017) 128:505–11. doi: 10.1016/j.clinph.2017.01.013
8. Salvalaggio A, Coraci D, Cacciavillani M, Obici L, Mazzeo A, Luigetti M, et al. Nerve ultrasound in hereditary transthyretin amyloidosis: red flags and possible progression biomarkers. *J Neurol*. (2020) 268:189–98. doi: 10.1007/s00415-020-10127-8

9. Tan C, Arumugam T, Razali SNO, Yahya MA, Goh K, Shahrizaila N. Nerve ultrasound can distinguish chronic inflammatory demyelinating polyneuropathy from demyelinating diabetic sensorimotor polyneuropathy. *J Clin Neurosci*. (2018) 57:198–201. doi: 10.1016/j.jocn.2018.08.031
10. Goedee HS, Herraets I, Visser LH, Franssen H, van Asseldonk JH, van der Pol WL, et al. Nerve ultrasound can identify treatment-responsive chronic neuropathies without electrodiagnostic features of demyelination. *Muscle Nerve*. (2019) 60:415–9. doi: 10.1002/mus.26629
11. Herraets IJT, Goedee HS, Telleman JA, van Eijk RPA, van Asseldonk JT, Visser LH, et al. Nerve ultrasound improves detection of treatment-responsive chronic inflammatory neuropathies. *Neurology*. (2020) 94:e1470–9. doi: 10.1212/WNL.0000000000008978
12. Puma A, Azulay N, Grecu N, Suply C, Panicucci E, Cambieri C, et al. Comparison of high-frequency and ultrahigh-frequency probes in chronic inflammatory demyelinating polyneuropathy. *J Neurol*. (2019) 266:2277–85. doi: 10.1007/s00415-019-09392-z
13. Sugimoto T, Ochi K, Hosomi N, Mukai T, Ueno H, Takahashi T, et al. Ultrasonographic reference sizes of the median and ulnar nerves and the cervical nerve roots in healthy japanese adults. *Ultrasound Med Biol*. (2013) 39:1560–70. doi: 10.1016/j.ultrasmedbio.2013.03.031
14. Allen JA, Ney J, Lewis RA. Electrodiagnostic errors contribute to chronic inflammatory demyelinating polyneuropathy misdiagnosis. *Muscle Nerve*. (2017) 57:542–9. doi: 10.1002/mus.25997
15. Kollmer J, Hund E, Hornung B, Hegenbart U, Schönland SO, Kimmich C, et al. *In vivo* detection of nerve injury in familial amyloid polyneuropathy by magnetic resonance neurography. *Brain*. (2015) 138:549–62. doi: 10.1093/brain/awu344
16. Leonardi L, Vanoli F, Fionda L, Loreti S, Garibaldi M, Morino S, et al. Nerve ultrasonography findings as possible pitfall in differential diagnosis between hereditary transthyretin amyloidosis with polyneuropathy and chronic inflammatory demyelinating polyneuropathy. *Neurol Sci*. (2020) 41:3775–8. doi: 10.1007/s10072-020-04717-7
17. Grimm A, Rasenack M, Athanasopoulou IM, Dammeier NM, Lipski C, Wolking S, et al. The modified ultrasound pattern sum score mUPSS as additional diagnostic tool for genetically distinct hereditary neuropathies. *J Neurol*. (2016) 263:221–30. doi: 10.1007/s00415-015-7953-7
18. Zaidman CM, Al-Lozi M, Pestronk A. Peripheral nerve size in normals and patients with polyneuropathy: an ultrasound study. *Muscle Nerve*. (2009) 40:960–6. doi: 10.1002/mus.21431
19. Niu J, Cui L, Liu M. Multiple sites ultrasonography of peripheral nerves in differentiating charcot–marie–tooth type 1A from chronic inflammatory demyelinating polyradiculoneuropathy. *Front Neurol*. (2017) 8:181. doi: 10.3389/fneur.2017.00181
20. Meng L, Lyu H, Zhang W, Liu J, Wang Z, Yuan Y. Hereditary transthyretin amyloidosis in eight chinese families. *Chinese Med J-Peking*. (2015) 128:2902–5. doi: 10.4103/0366-6999.168048
21. Sueyoshi T, Ueda M, Jono H, Tasaki M, Murata S, Horibata Y, et al. Transthyretin-derived amyloidosis in musculoskeletal systems. *Amyloid*. (2011) 18(Suppl 1)163–5. doi: 10.3109/13506129.2011.574354061
22. Luan X, Zheng R, Chen B, Yuan Y. Childhood chronic inflammatory demyelinating polyneuropathy with nonuniform pathologic features. *Pediatr Neurol*. (2010) 43:103–9. doi: 10.1016/j.pediatrneurol.2010.04.001
23. Ikeda S, Koike H, Nishi R, Kawagashira Y, Iijima M, Katsuno M, et al. Clinicopathological characteristics of subtypes of chronic inflammatory demyelinating polyradiculoneuropathy. *J Neurol Neurosurg Psychiatr*. (2019) 90:988–96. doi: 10.1136/jnnp-2019-320741

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 26 February 2021
doi: 10.3389/fneur.2021.630387

Fast Open-Source Toolkit for Water T2 Mapping in the Presence of Fat From Multi-Echo Spin-Echo Acquisitions for Muscle MRI

Francesco Santini[1,2], Xeni Deligianni[1,2], Matteo Paoletti[3], Francesca Solazzo[3], Matthias Weigel[1,4,5], Paulo Loureiro de Sousa[6], Oliver Bieri[1,2], Mauro Monforte[7], Enzo Ricci[7,8], Giorgio Tasca[7], Anna Pichiecchio[3,9†] and Niels Bergsland[10,11†]*

[1] Division of Radiological Physics, Department of Radiology, University Hospital of Basel, Basel, Switzerland, [2] Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland, [3] Advanced Imaging and Radiomics Center, Neuroradiology Department, IRCCS Mondino Foundation, Pavia, Italy, [4] Translational Imaging in Neurology (ThINk) Basel, Department of Biomedical Engineering, University Hospital Basel and University of Basel, Allschwil, Switzerland, [5] Neurologic Clinic and Policlinic, Departments of Medicine, Clinical Research and Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland, [6] ICube, Université de Strasbourg, Centre National de la Recherche Scientifique (CNRS), Strasbourg, France, [7] Unità Operativa Complessa di Neurologia, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy, [8] Dipartimento di Neuroscienze, Istituto di Neurologia, Università Cattolica del Sacro Cuore, Rome, Italy, [9] Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy, [10] Department of Neurology, Buffalo Neuroimaging Analysis Center, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, The State University of New York, Buffalo, NY, United States, [11] Fondazione Don Carlo Gnocchi Onlus (IRCCS), Milan, Italy

Imaging has become a valuable tool in the assessment of neuromuscular diseases, and, specifically, quantitative MR imaging provides robust biomarkers for the monitoring of disease progression. Quantitative evaluation of fat infiltration and quantification of the T2 values of the muscular tissue's water component (wT2) are two of the most essential indicators currently used. As each voxel of the image can contain both water and fat, a two-component model for the estimation of wT2 must be used. In this work, we present a fast method for reconstructing wT2 maps obtained from conventional multi-echo spin-echo (MESE) acquisitions and released as Free Open Source Software. The proposed software is capable of fast reconstruction thanks to extended phase graphs (EPG) simulations and dictionary matching implemented on a general-purpose graphic processing unit. The program can also perform more conventional biexponential least-squares fitting of the data and incorporate information from an external water-fat acquisition to increase the accuracy of the results. The method was applied to the scans of four healthy volunteers and five subjects suffering from facioscapulohumeral muscular dystrophy (FSHD). Conventional multi-slice MESE acquisitions were performed with 17 echoes, and additionally, a 6-echo multi-echo gradient-echo (MEGE) sequence was used for an independent fat fraction calculation. The proposed reconstruction software was applied on the full datasets, and additionally to reduced number of echoes, respectively, to 8, 5, and 3, using EPG and biexponential least-squares fitting, with and without incorporating information from the MEGE acquisition. The incorporation of external fat fraction maps increased the robustness of the fitting with a reduced number of echoes per datasets, whereas with unconstrained fitting, the total of 17 echoes was necessary to retain an independence of wT2 from the level of fat infiltration. In conclusion, the

OPEN ACCESS

Edited by:
Itamar Ronen, Leiden University Medical Center, Netherlands

Reviewed by:
Dominik Weidlich, Technical University of Munich, Germany
Doris Leung, Kennedy Krieger Institute, United States

***Correspondence:**
Francesco Santini
francesco.santini@unibas.ch

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Applied Neuroimaging, a section of the journal Frontiers in Neurology

Received: *17 November 2020*
Accepted: *05 February 2021*
Published: *26 February 2021*

Citation:
Santini F, Deligianni X, Paoletti M, Solazzo F, Weigel M, de Sousa PL, Bieri O, Monforte M, Ricci E, Tasca G, Pichiecchio A and Bergsland N (2021) Fast Open-Source Toolkit for Water T2 Mapping in the Presence of Fat From Multi-Echo Spin-Echo Acquisitions for Muscle MRI. Front. Neurol. 12:630387. doi: 10.3389/fneur.2021.630387

proposed software can successfully be used to calculate wT2 maps from conventional MESE acquisition, allowing the usage of an optimized protocol with similar precision and accuracy as a 17-echo acquisition. As it is freely released to the community, it can be used as a reference for more extensive cohort studies.

Keywords: MRI, neuromuscular diseases, relaxometry, free open source software, water T2 relaxation time, fat water imaging

INTRODUCTION

Neuromuscular disorders encompass genetic and acquired diseases of lower motor neurons, peripheral nerves, neuromuscular junction, or skeletal muscle, generally causing different degrees of motor impairment in the affected patients. In particular, muscular dystrophies are hereditary degenerative disorders of skeletal muscles, causing progressive replacement of muscle tissue by fat. This happens through disparate pathological processes and molecular mechanisms that are, at least to some extent, disease-specific and related to the peculiar genetic defect that characterizes each of them (1).

However, some of these broad pathological processes are shared by most muscular dystrophies. They can be followed on muscle imaging as an early phase of muscle damage and intramuscular edema, often corresponding to inflammatory/necrotic changes (2, 3), and subsequent stages characterized by progressive deposition of fat and connective tissue (4–6). Different pathological processes are generally present simultaneously in the same patient or even in the same muscle group.

The good soft tissue characterization capabilities of MRI can be exploited to quantify the status of these ongoing processes in the muscle; for example, global T2 contrast can be used to highlight edema, and fat/water separation methods can show fat infiltration.

When moving toward quantitative imaging, global T2, extracted by a monoexponential fitting is a sensitive disease indicator when fat infiltration is not present. However, in neuromuscular diseases where adipocytes significantly replaced muscular tissue, it highly correlates with the fat content of the muscle. In this case, it is therefore rather an indicator of long-term changes in the musculature, albeit an indirect one with respect to fat fraction quantification; conversely, the T2 relaxation of the water component of the tissue (wT2) well correlates with acute "disease activity" (7).

Various acquisition methods have been proposed to quantify wT2 independently of fat infiltration (8–15). In current clinical practice, multi-echo spin-echo (MESE) sequences are typically used, and various types of exponential fitting (biexponential, triexponential) are used for T2 calculation (16, 17). However, these methods are sensitive to multiple confounding factors, such as B1 inhomogeneities.

Marty et al. (14) presented a method based on an extended phase graph (EPG) fitting that addresses many of these issues. Besides, in contrast to other methods such as the ones proposed by Klupp et al. (12), Sousa et al. (11), or Koolstra et al. (13), the EPG-based approach has the advantage of using a conventional spin-echo sequence for the quantification, which is broadly available and therefore does not require any particular sequence modification or hardware. In recent years, this method has been extensively used for wT2 quantification in several studies (18–22). Although there are small differences in the exact protocol used depending on the MR scanner vendor, the differences in the implementation of the fitting process can be more substantial.

As a result of the EPG fitting, a relatively precise estimation of the fat fraction is also obtained. The resulting fat fractions approximate the results of a three-point Dixon acquisition, which has been proven to differentiate well between patients of various neuromuscular diseases and healthy subjects (16, 23). However, fat fraction estimation is not the primary quantitative target of this method, and it can be an issue when accurate fat fraction maps are needed, as weighting factors [such as magnetization transfer effects in multi-slice MESE acquisitions (24–26)] can bias the estimation. An alternative approach to obtain this information is to use a dedicated sequence (usually based on small flip-angle multi-echo gradient-echo imaging) for a purely proton-density-weighted fat fraction estimation. This approach also allows using more accurate fat models that incorporate the complex chemical composition of fat (27). In practice, a typical muscle MRI examination usually comprises a dedicated volume acquisition for single muscle identification (segmentation) and for accurate fat fraction acquisition.

In this work, we build upon this concept to present a fast and open software for wT2 fitting, which can be used as a reference implementation for reproducibility studies. The postprocessing performance is optimized by the usage of GPU processing and the creation of a dictionary (of adjustable size) of simulated signals incorporating slice profile information (14, 28). Additionally, this work further extends the initial concept as it implements multiple fitting methods for the wT2 estimation, and it can incorporate the information coming from a separate fat/water acquisition (of arbitrary resolution and field of view) to constrain the fitting. With this constraint, the possibility of reducing the number of acquired echoes for the fitting (thus reducing the scan time) is also analyzed.

MATERIALS AND METHODS

Software Implementation

A stand-alone, command-line application was developed in Python, using standard mathematical extension libraries (NumPy, SciPy) and, for hardware acceleration, the pycuda extension for interfacing with the general-purpose graphical processing unit.

The core of the application consists of a fast, GPU-accelerated implementation of signal simulation based on the extended phase graph concept (29, 30). As the method is intended for multi-echo spin-echo acquisitions, a Carr-Purcell-Meiboom-Gill (CPMG) (31, 32) simulation is performed at every run of the program, adapting the timing and the number of echoes to the actual sequence parameters. The slice profile of the radiofrequency pulses is taken into account in the simulation, assuming hanning-windowed sinc pulses. The slice profile is calculated through the application of a Shinnar-LeRoux transform (33) and the slice width of the refocusing pulse is assumed to be a factor of 1.2 larger than the excitation pulse, as per characteristics of the pulse sequence of the used MR scanner. For different vendors and acquisition protocols, the user can provide custom slice profiles and refocusing width factors as parameters to the program. The simulation logic is directly written in GPU-specific C++ language for maximum efficiency. Thanks to the high parallelization of the GPU tasks, signals corresponding to multiple combinations of wT2, fat fraction, and B1 inhomogeneity factors can be simulated concurrently and placed in a dictionary, whose size can be chosen by the user. The fat model is assumed to have a single spectral peak with a fixed T2 value, which can either be given *a priori* or estimated from the input data. T1 values for both water and fat are assumed to be constant (1,400 and 365 ms, respectively), as in (14).

Subsequently, the time course of each voxel in the input series is compared to each entry in the dictionary by using the correlation metric. The parallelization of the GPU is again leveraged in this step, as the correlation operation is equivalent to the following matrix multiplication:

$$\mathbf{C} = \mathbf{S} \times \mathbf{D}$$

where **S** is the signal matrix, having a separate voxel on each row and time in the column direction; **D** is the dictionary matrix, having time in the row direction and dictionary entries in the row directions; **C** is the result matrix holding the signal correlations with the dictionary. After multiplication, the index of the maximum value of each column identifies the parameter combination that best fits the measured signal.

The above procedure describes the unconstrained fitting. When an externally derived fat fraction map is also provided, it is aligned (based on the slice orientation and position) with each slice of the original spin-echo data. The information is then taken into account by restricting the maximum search to the given fat fraction for each voxel.

In addition to EPG fitting, the software can also perform double-exponential fitting. This fitting can neither correct for B1 inhomogeneity nor slice profile and is meant to be used when the EPG fitting is unstable or the true slice profile is not known.

The code is released under a free software license (GNU General Public License v3) at the website: https://www.github.com/fsantini/MyoQMRI.

Acquisition Protocol

A conventional 2D multi-slice multi-echo spin-echo (MESE) acquisition protocol was prepared for a commercial whole-body 3T MR scanner (MAGNETOM Skyra, Siemens Healthcare, Erlangen, Germany) equipped with an 18-channel body array coil and integrated spine coil with the following acquisition parameters: number of echoes 17, number of slices 7, TR 4,100 ms, first TE and echo spacing 10.9 ms, bandwidth 250 Hz/px, matrix size 192 × 384, resolution 1.2 × 1.2 mm^2, slice thickness 10 mm, gap between slices 30 mm.

For fat fraction quantification, a 3D multi-echo gradient-echo (MEGE) acquisition using a custom sequence was prepared with the following parameters: number of echoes 6, TR 35 ms, first TE/echo spacing 1.7/1.5 ms, flip angle 7°, bandwidth 1,050 Hz/px, matrix size 396 × 432 × 52, resolution 1.0 × 1.0 × 5.0 mm^3. The sequence had a monopolar readout with interleaved echo spacing (even and odd echoes acquired in subsequent repetitions). The acquired images were postprocessed with the publicly-available algorithm FattyRiot (27) to obtain the fat fraction maps.

Human Experiments

The acquisitions described above were performed on the thighs of 5 (three male, median age 46 y, range 29–61 y) subjects with diagnosed facioscapulohumeral muscular dystrophy (FSHD), which is a particular form of muscular dystrophy characterized by progressive muscle wasting and fatty replacement often preceded by muscle edema on MRI (34), and of four subjects without a history of neuromuscular diseases (two male, median age 54.5 y, range 26–72 y). The acquisitions were performed according to the local ethics regulations and informed consent was obtained from the participants.

Data Analysis

The computer used for the postprocessing is a current mid-range personal computer (Ryzen 2600, Advanced Micro Devices, Santa Clara, CA, equipped with 32GB of RAM and a GeForce 1060 GPU, NVIDIA corporation, Santa Clara, CA).

The proposed algorithm was applied to the MESE acquisition after generation of a dictionary containing 60 linearly spaced values for wT2 (range 20–80 ms), 20 values for B1 factor (range 40–140%), and 101 values for the fat fraction (range 0–100%), for a total of 121,200 parameter combinations. The fat T2 was estimated from a subsample of the subjects (through a single-component EPG fitting in regions of subcutaneous fat) and subsequently assumed constant at 151 ms. Maps derived from EPG and double exponential fitting, with and without the constraint of the external proton-density-weighted fat fraction as calculated from the MEGE acquisition, were produced (an overview of the assumed and fitted parameters for each method is given in **Table 1**). The fitting was repeated by discarding later spin echoes (and only retaining 8, 5, or 3) to evaluate the robustness of the algorithm with respect to the number of echoes.

Regions of interest (ROIs) were manually drawn by one reader bilaterally on every MESE slice over the cross section of the following muscles: Vastus Lateralis (VM), Vastus Medialis (VM), Rectus Femoris (RF).

The average and standard deviation for wT2 values and fat fraction were extracted, and the following indicators were calculated:

TABLE 1 | Overview of the presented fitting methods.

Method	Constrained variables	Fitted parameters
EPG unconstrained	Water T1, Fat T1, Fat T2, Slice profile	wT2, Fat Fraction, B1
EPG constrained	Water T1, Fat T1, Fat T2, Slice profile, Fat Fraction (voxelwise from MEGE)	wT2, B1
Double exponential unconstrained	Fat T2	wT2, Fat Fraction
Double exponential constrained	Fat T2, Fat Fraction (voxelwise from MEGE)	wT2

- Average error between fat fraction calculated from MESE and the one deriving from MEGE (considered the "gold standard") - only for nonconstrained reconstructions, calculated as:

$$ERR_{ff} = \underset{pop,subj,roi}{\text{mean}}\,(FF_{MESE}) - \underset{pop,subj,roi}{\text{mean}}\,(FF_{MEGE}),$$

  where $\underset{pop,subj,roi}{\text{mean}}$ represents the averaging operation over the whole population, over all ROIs of each subject, and over the voxels of each ROI.
- Pooled standard deviation across the ROIs—this indicator is similar to an "average of the standard deviations," and it indicates the variability of the fitted wT2 values over a small spatial region; it is related to noise or tissue inhomogeneities due to disease activity and is calculated as follows:

$$SD_{pooled} = \sqrt{\underset{pop,subj}{\text{mean}}\left(\underset{roi}{\text{sd}^2}\,(T_2)\right)},$$

  where $\underset{roi}{\text{sd}^2}$ represents the squared standard deviation over each ROI.
- Standard deviation of the averages over the ROIs (hereby termed "intrasubject standard deviation") - this indicator relates to the variability over larger areas, for example arising from field inhomogeneity, and is calculated as follows:

$$SD_{intrasubject} = \sqrt{\underset{pop}{\text{mean}}\left(\underset{subj}{\text{sd}^2}\left(\underset{roi}{\text{mean}}\,(T_2)\right)\right)}$$

  This indicator was only calculated in the volunteer cohort, because patients might have variability in the wT2 values due to their pathology.
- Global average wT2 for patients and control groups, calculated as follows:

$$\text{mean}_{T2,(patients|controls)} = \underset{patients|controls}{\text{mean}}\left(\underset{subj,roi}{\text{mean}}\,(T_2)\right)$$

$$\text{sd}_{T2,(patients|controls)} = \underset{patients|controls}{\text{sd}}\left(\underset{subj,roi}{\text{mean}}\,(T_2)\right)$$

Statistical analysis was performed with R (35). To quantify the influence of the fat fraction on the fitting of wT2, the correlation between the average fat fraction and the average wT2 for each ROI was calculated. For the calculation of the correlation coefficients the log-transformed fat fractions were used to compensate for the skewness of the fat fraction distribution. However, the visualization was done on the original fat fraction axis (0–100%) for ease of interpretation.

In addition, correlation of repeated measurements (rmcorr function) instead of the standard Pearson coefficient was used to compensate for dependence of measurements of the same subject. In this case the subject is introduced as a variable and both wT2 and fat fraction are considered as measures.

Validation

In order to evaluate the absolute accuracy of the fitting method, the same data was analyzed using the EPG wT2 fitting procedure of the QMRTools software package (36). The slice profile was adapted to match the one used in the current acquisitions and the fitting was applied to the full 17-echo dataset.

The same ROIs as in the previous analysis were considered, and the average wT2 in each ROI was calculated and compared to the corresponding ROI of our method. Average and standard deviation of the errors were calculated, and the agreement was visualized in a Bland-Altman plot.

RESULTS

The software fitted the MESE datasets with EPG dictionary matching in an average time of 2 s/slice, plus 12 s per dataset for the generation of the dictionary (operation performed at every run of the program, which was faster than caching the results to disk). The double exponential fitting was not GPU-optimized and took 390 s (6 m 30 s) per slice.

Exemplary outputs from an unconstrained 17-echo reconstruction for a patient and a volunteer are shown in **Figure 1**. No noticeable artifacts are visible in the wT2 and fat fraction maps, whereas the B1 maps show some inconsistencies, mostly located in the fascia, thus remaining mostly masked in the other maps.

WT2 maps had a homogeneous appearance with EPG matching both for 17 and 8 echoes (**Figure 2**). However, the unconstrained reconstruction showed a noticeable difference in quantitative values in areas with heavy fat infiltration, whereas the fat-fraction-constrained reconstructions appeared similar between 17, 8, and 5 considered echoes. The visual quality of the map was insufficient in any combination when only three echoes were used for the reconstruction (**Figure 2**). Quantitatively, it can be observed that a reduced number of echoes in the unconstrained fitting has an effect on wT2 fitting when fat infiltration is present, resulting in a highly significant positive correlation ($p < 0.001$) for EPG fitting with 5 and 8 echoes (correlation coefficient $r = 0.73$ and 0.61, respectively, **Figure 3**).

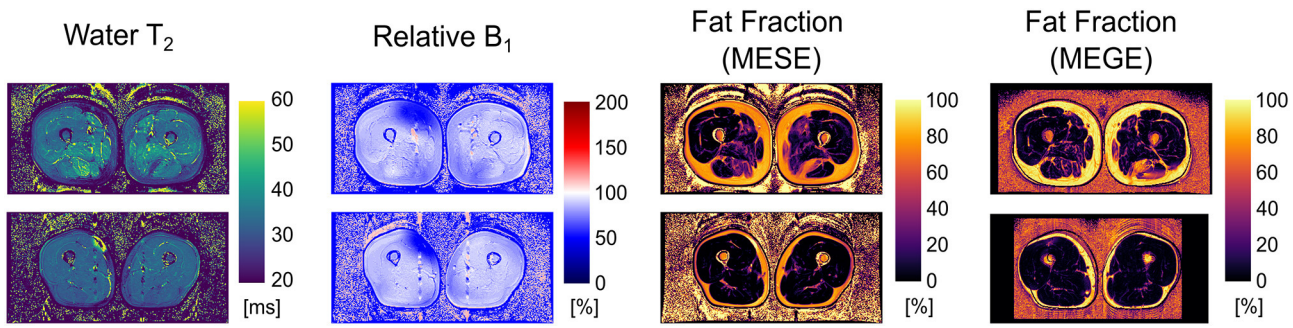


FIGURE 1 | Parameter maps derived from an unconstrained 17-echo EPG dictionary matching for an FSHD patient (**top row**) and a healthy volunteer (**bottom row**). The rightmost panel is the registered fat fraction map deriving from a multi echo gradient echo acquisition (MEGE) for reference.

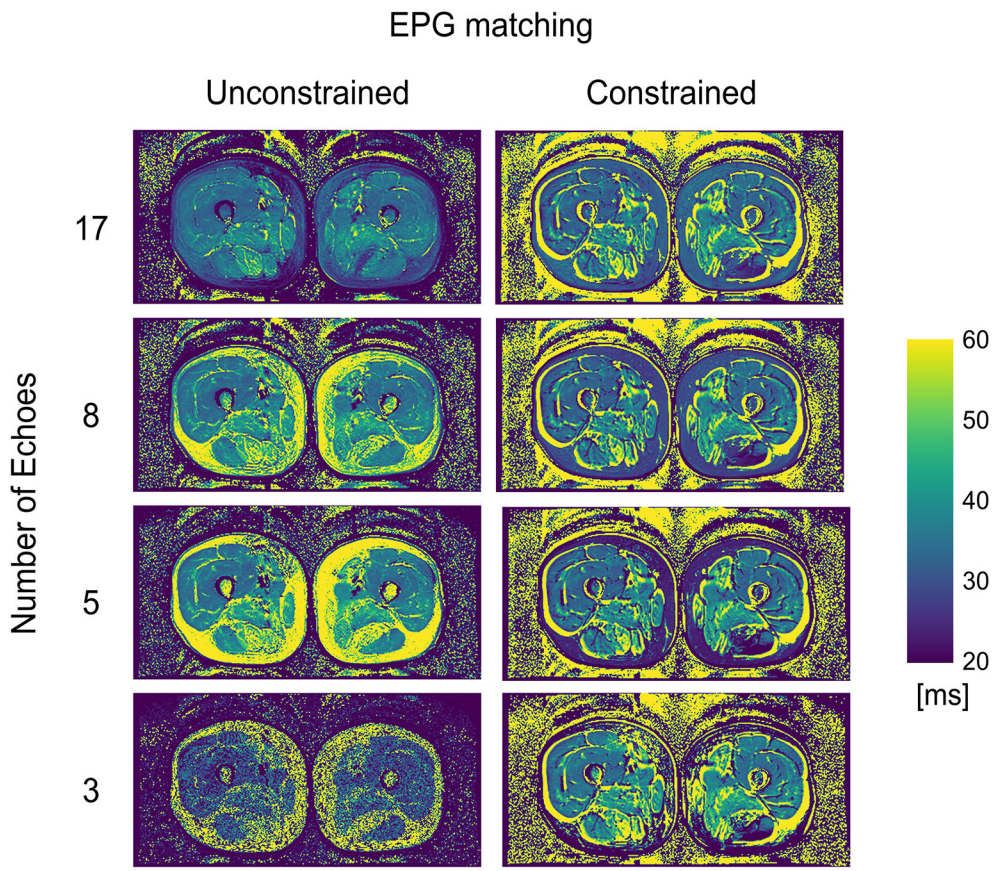


FIGURE 2 | wT2 maps for an FSHD patient using a varying number of acquired echoes by EPG simulation with and without an external fat fraction map as a reconstruction constraint.

Conversely, the correlation was significantly negative ($r = -0.71$, $p < 0.001$) for the unconstrained fitting with 17 echoes, and for the constrained fitting of 5 and 8, but not 17 echoes ($r = -0.60$, -0.41, -0.19; $p < 0.001$, $p = 0.005$, 0.2, respectively).

The double exponential fitting showed decreased wT2 with any number of echoes when unconstrained fitting was used, in addition to artifacts arising from B1 inhomogeneities (see **Figure 4** and **Table 2**). Correlation between wT2 and the fat fraction was always negative and significant ($p = 0.02$ for the constrained 8 and 17 echoes, $p < 0.001$ otherwise) in all cases except, notably, the unconstrained fitting of five echoes ($p = 0.29$), with negative correlation coefficients ranging from $r = -0.34$ for the constrained 8-echo reconstruction to $r = -0.62$ for the unconstrained 8-echo reconstruction (**Figure 3**).

According to the quantitative quality metrics (**Table 2**), considering a higher number of echoes improved both the noise

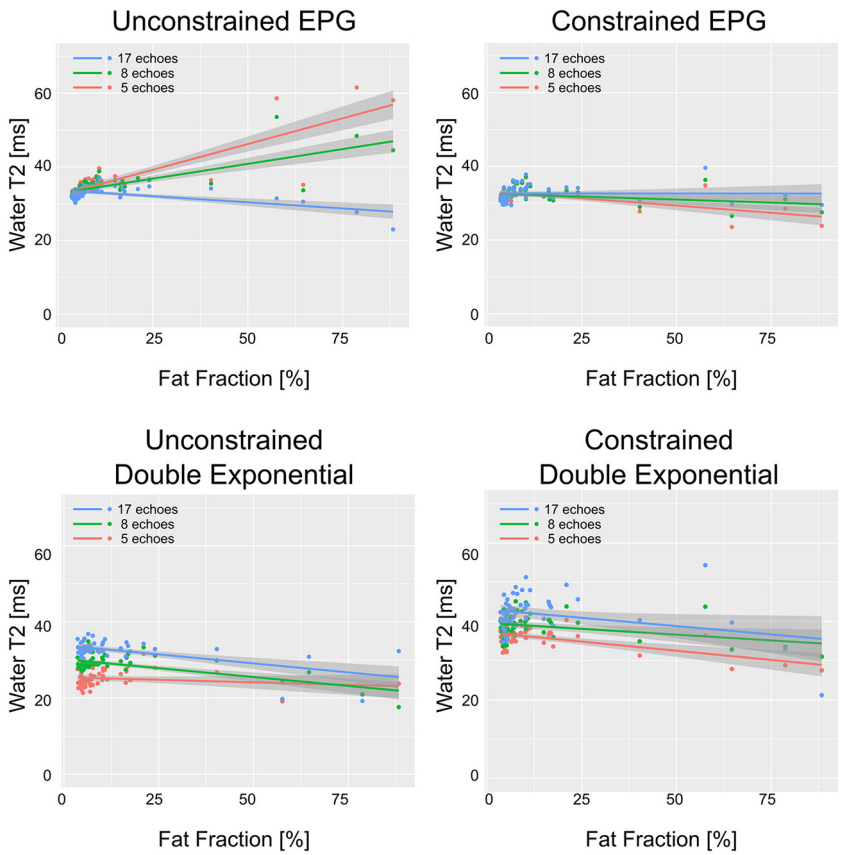


FIGURE 3 | wT2 values over each ROI vs. fat fraction in the same ROI, for all ROIs drawn in patients and healthy volunteers. The solid lines represent the linear regression with shaded gray areas indicating the 95% confidence intervals.

(pooled standard deviation) and the homogeneity across the ROIs for each subject (intrasubject variability). For a reduced number of echoes, the constrained reconstruction retained good homogeneity down to five considered echoes for the EPG reconstruction (3.1 ms for the constrained case vs. 6.7 for the unconstrained). The double exponential fitting showed good homogeneity (intrasubject standard deviation ranging from 2.7 to 4.6 milliseconds), but more noise (pooled standard deviation ranging from 5.9 to 10.5 ms).

The constrained reconstruction resulted in a generally higher pooled standard deviation due to the artifacts introduced by the misregistration of the images.

The accuracy of the wT2 values was close to the QMRTools implementation, showing an average difference of 0.9 ms (5.2 ms standard deviation, **Figure 5**).

Concerning the fat fraction accuracy, although it is not the primary focus of this method, the EPG matching resulted in consistent underestimation compared to the FattyRiot algorithm, with a larger number of echoes providing results which were closer in average to the gold standard (from −4.2 percentage points (p.p.) for 17 echoes to -7.8 p.p. for eight echoes). The double exponential fitting provided an overestimation ranging from +3.2 p.p. for 17 echoes to +8.6 p.p. for five echoes.

DISCUSSION

In this work, we presented a software application that can quickly and reliably calculate wT2 maps in the presence of spectrally inhomogeneous voxels containing both non-fatty tissue and lipids, while at the same time estimating fat fraction from conventional multi-echo spin-echo acquisitions. This work follows the concept introduced by Marty et al. (14). Still, it additionally provides the possibilities of performing double exponential (and, optionally, single exponential) fitting and, more interestingly, of incorporating external fat fraction information to improve the accuracy of the fitting. There are no specific hardware requirements for this program; however, for better performance, a reasonable Cuda-compatible graphic card should be used. Consumer-grade GPUs deliver good performance, but more extensive dictionaries than the one tested in this current setup require increased memory on the device. The code is released with a free open source license and, in contrast to existing available implementations (36), this software package is exclusively based on free software and is platform-independent. This implementation can thus be considered ready to be widely used in a clinical research context and as a reference by other researchers.

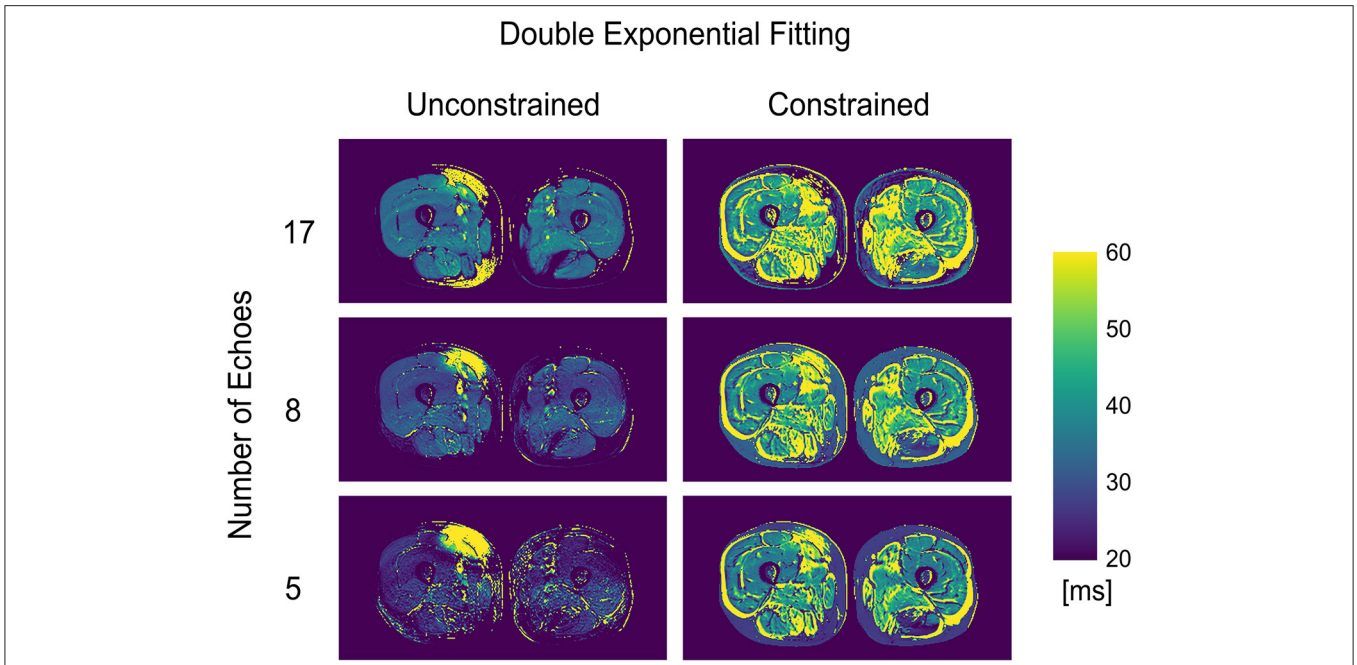


FIGURE 4 | wT2 maps for an FSHD patient using a varying number of acquired echoes by double exponential fitting with and without an external fat fraction map as a reconstruction constraint.

TABLE 2 | Summary of results for the different fitting methods, with and without an external proton-density-weighted fat fraction (FF) constraint.

Method	FF Constraint	# echoes	wT2 (ms)							FF (%)		Global FF error (p.p.)
			Volunteers				Patients			Vol.	Pat.	
			Average	Global SD	Pooled SD	Intrasubject SD	Average	Global SD	Pooled SD			
EPG	No	3	28.6	1.3	8.1	1.0	31.4	6.2	12	17.3	22.5	6.4
		5	34.2	1.8	5.2	1.1	37.7	7.7	8.8	2.9	8.1	−8.0
		8	33.8	1.8	4.4	1.0	36.1	4.9	7.6	3.2	9.8	−7.0
		17	33.0	1.3	3.3	0.9	32.7	2.5	4.1	4.6	12.8	−4.7
	Yes	3	34.5	1.5	5.9	1.5	34.5	3.6	8.9	6.4*	17.8*	
		5	32.3	1.3	4.9	1.1	32.2	3.2	6.4			
		8	31.9	1.4	4.9	1.3	32.4	2.6	5.9			
		17	32.1	1.6	5.1	1.4	33.1	2.5	6.2			
Double exponential	No	5	25.3	2.9	8.3	3.1	24.7	2.5	10.1	17.9	23.9	8.6
		8	29.4	1.7	5.8	1.8	28.4	3.1	6.7	15.1	21.9	6.2
		17	33.2	1.6	4.6	1.6	31.7	3.7	9.5	11.9	19.0	3.2
	Yes	5	36.1	2.1	6.9	2.3	35.4	3.6	8.3	6.4*	17.8*	
		8	38.6	2.5	7.7	2.7	38.4	3.7	9.3			
		17	41.3	3.5	10.4	3.5	41.2	6.2	14			

The pooled standard deviation (SD) is associated with image noise, whereas the intrasubject SD is associated with homogeneity in different anatomical regions. The FF error with respect to the multi-echo gradient-echo acquisition (MEGE) is given in percentage points (p.p.). Values marked with an asterisk (*) are not fitted by the algorithm but are derived from the MEGE acquisition.

In the comparison with an existing implementation, the absolute wT2 values obtained by this method are close, but not identical, to the ones obtained from QMRTools, although both methods are based on the same conceptual framework. One explanation could be the usage of different metrics while performing the dictionary matching. This finding highlights the necessity of having consistent acquisition and data processing pipelines, and the necessity of characterizing a quantitative method primarily in terms of reproducibility and precision. It is generally true also in other fields of quantitative MRI, that

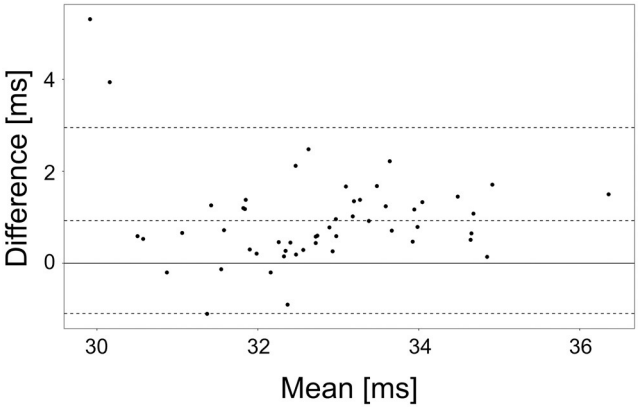


FIGURE 5 | Bland Altman plot of the agreement of the wT2 values obtained by the QMRTools software package and the presented software, for an unconstrained 17-echo EPG reconstruction. The bias and the 95% confidence intervals are depicted in the plot.

the absolute values obtained are method-dependent and thus comparisons need to be carefully considered (37, 38).

During the optimization of the EPG matching algorithm, it appeared clear that an accurate slice profile was very important to obtain absolute values of wT2 close to the literature. Small changes to the profile, or selecting a too large refocusing slice width factor, could introduce a bias of a few milliseconds in the estimated values. The intersubject and intrasubject variabilities, however, minimally changed, so the user should be advised to obtain the exact sequence characteristics and to use a coherent parameter set when reconstructing multiple datasets.

The introduction of the constrained reconstruction appears beneficial for the fitting of wT2, providing reliable and homogeneous results, not correlated with the amount of infiltration of the muscle, with considering as few as five acquired echoes. This is a valuable result in light of providing better spatial coverage by the MESE sequence: reducing the number of acquired echoes allows exciting multiple interleaved slices in a simple repetition time and thus lessen the interslice gaps. However, the alignment of the fat fraction map can introduce some artifacts if patient motion occurs between the two acquisitions. In the current implementation, no image registration is performed. The alignment of the MESE and MEGE datasets is currently only performed based on the orientation information provided in the image headers. Image registration is, however, implementable using free python libraries and could be added if needed.

The EPG matching, as expected, could account for the B1 inhomogeneities and therefore lead to lower artifacts in the areas where flip angles deviate from the nominal value in comparison to double-exponential fitting.

In general, unconstrained double exponential fitting with reduced echoes resulted in underestimating wT2 values relative to other methods with unconstrained reconstruction. On the other hand, thanks to its fewer degrees of freedom in the fitting, it showed remarkable homogeneity across the various ROIs even with few echoes, suggesting that it might still be considered a robust and straightforward approach when no other methods are available. The constrained double exponential reconstruction, conversely, produced results farthest from the expected values (12, 14). An analysis of the data shows that the unconstrained double exponential fitting consistently overestimates the fat fraction. The possible explanation is that the longer exponential decay due to the stimulated echoes gets assigned to the fat component during the fitting process in the unconstrained case, whereas fixing the fat fraction in the constrained reconstruction produces a longer apparent exponential decay constant in what is practically a monoexponential fitting of the residuals.

Generally, most of the obtained wT2 values negatively correlate with fat fraction, which is in line with the previous findings (15, 39), with the exception of the EPG matching of the reduced echoes, showing a positive correlation, suggesting a failure to separate the wT2 from fat.

As an overall comparison, both fitting methods (double exponential fitting and EPG matching) have advantages and disadvantages. EPG matching appears accurate and precise even when a lower number of echoes is used, especially when paired with an external fat fraction constraint, and it has a higher insensitivity to B1 inhomogeneity, but it requires precise knowledge of the acquisition parameters to obtain unbiased values. Double exponential fitting fails in regions of B1 inhomogeneity, its accuracy is poor when few echoes are used, and the external fat fraction constraints introduces a bias in the obtained values. However, it requires very little knowledge of the acquisition parameters, and it can therefore be chosen when the sequence characteristics are unknown. When the full 17-echo acquisitions are used, the two fitting methods have similar characteristics when averaged across the muscles; however, the sensitivity of the double exponential fitting to B1 inhomogeneity might mask local intramuscular changes in patients with neuromuscular disorders.

Although not explored in the present validation for the results to be more comparable, the program assumes a constant value for fat T2. Still, it gives the possibility of indicating it as a runtime parameter or estimating it from the image itself. Similarly, T1 values for water and fat are fixed, although this can be assumed to have a small effect on the final result (14). Other effects like j-coupling are also not explicitly introduced in the model but rather incorporated in the assumed value of fat T2, which is dependent on the characteristics of the MESE sequence (40).

The implementation has some further limitations regarding the accuracy of the results. Specifically, Keene et al. demonstrated that this method can be improved by introducing a correction for the chemical shift in the slice profile and a better estimation for fat T2 (15). These corrections are relatively newly introduced and not routinely used in the current studies, and thus not currently implemented. However, the chemical shift correction requires knowledge of the actual implementation of the pulse sequence that might be difficult to obtain in a clinical setting. It would therefore result in potential loss of generalizability. Similarly, a multi-peak spectral model for the fat could be incorporated into the algorithm, which could improve the results' accuracy. As this method is based on multi spin echoes, the multiple spectral

peaks do not result in different chemical shifts as in gradient-echo images. The effect would only be seen in the different T2 values of the fat components; therefore, this functionality is not currently implemented, in line with existing spin-echo-based methods for muscle imaging (14, 15, 17, 23).

One limitation of this work is the relatively small number of datasets. However, both healthy and different diseased subjects were included. While a larger subject cohort would be of interest, one of the goals of this study was to offer the tools for such studies and not to draw conclusions on wT2 measurements of dystrophic muscles. For this goal, clinical studies that focus on more homogenous patient cohorts (grouping, for example, the different stages of neuromuscular disease, where similar pathological changes are expected) are required.

The validation of this work was performed by comparing the proposed fitting with an existing implementation of the same concept on the same data, thus lacking an external reference standard for the values obtained. While such an external reference could be desirable, the scope of this work was not to assess the validity of MESE acquisitions for the estimation of wT2, but only to evaluate the efficacy of the proposed implementation. The availability of an independent fitting procedure allowed a direct comparison of these characteristics without other sources of error deriving from the physics of different acquisition methods.

In conclusion, in this work, we have presented a fast and open implementation of an algorithm for the T2 mapping of the water component of muscle tissues in the presence of fat, based on conventional multi-echo spin-echo sequences, capable of incorporating prior knowledge of the fat fraction. Thanks to the additional information obtained from the multi-echo gradient-echo images, a reduced number of echoes can be used for the spin-echo acquisition with while retaining similar inter- and intrasubject variability and similar absolute values, when and EPG model is used. The EPG matching method is in general to be preferred to a double exponential fitting, provided that the characteristics of the MR sequence (especially in terms of RF pulses) are sufficiently known.

DATA AVAILABILITY STATEMENT

The data analyzed in this study is subject to the following licenses/restrictions: Data included in this work can be shared upon research agreement between the Institutions. Requests to access these datasets should be directed to Anna Pichiecchio, anna.pichiecchio@mondino.it.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comitato Etico Area Referente Pavia Fondazione IRCCS Policlinico San Matteo and Policlinico Gemelli's Ethics Committe. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FSa study conception, method and software development, data analysis, and manuscript drafting. XD data analysis, statistics, and manuscript drafting. MP, FSo, and AP data acquisition and patient management. MW and PS method development. MM, ER, and GT patient recruitment and clinical evaluations. OB, AP, and NB study conception and supervision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Swiss National Science Foundation (SNSF) grant no. 172876 and by the Italian Ministry of Health (RC 2017-2019 and RF-2016-02362914).

REFERENCES

1. Mercuri E, Muntoni F. Muscular dystrophies. *Lancet Lond Engl.* (2013) 381:845–60. doi: 10.1016/S0140-6736(12)61897-2
2. Tasca G, Pescatori M, Monforte M, Mirabella M, Iannaccone E, Frusciante R, et al. Different molecular signatures in magnetic resonance imaging-staged facioscapulohumeral muscular dystrophy muscles. *PLoS ONE.* (2012) 7:e38779. doi: 10.1371/journal.pone.0038779
3. Lassche S, Küsters B, Heerschap A, Schyns MVP, Ottenheijm CAC, Voermans NC, et al. Correlation between quantitative MRI and muscle histopathology in muscle biopsies from healthy controls and patients with IBM, FSHD and OPMD. *J Neuromuscul Dis.* (2020) 7:495–504. doi: 10.3233/JND-200543
4. Monforte M, Laschena F, Ottaviani P, Bagnato MR, Pichiecchio A, Tasca G, et al. Tracking muscle wasting and disease activity in facioscapulohumeral muscular dystrophy by qualitative longitudinal imaging. *J Cachexia Sarcopenia Muscle.* (2019) 10:1258–65. doi: 10.1002/jcsm.12473
5. Klingler W, Jurkat-Rott K, Lehmann-Horn F, Schleip R. The role of fibrosis in Duchenne muscular dystrophy. *Acta Myol Myopathies Cardiomyopathies.* (2012) 31:184–95.
6. Bonati U, Hafner P, Schädelin S, Schmid M, Naduvilekoot Devasia A, Schroeder J, et al. Quantitative muscle MRI: a powerful surrogate outcome measure in Duchenne muscular dystrophy. *Neuromuscul Disord.* (2015) 25:679–85. doi: 10.1016/j.nmd.2015.05.006
7. Carlier PG. Global T2 versus water T2 in NMR imaging of fatty infiltrated muscles: different methodology, different information and different implications. *Neuromuscul Disord.* (2014) 24:390–2. doi: 10.1016/j.nmd.2014.02.009
8. Janiczek RL, Gambarota G, Sinclair CDJ, Yousry TA, Thornton JS, Golay X, et al. Simultaneous T2 and lipid quantitation using IDEAL-CPMG. *Magn Reson Med.* (2011) 66:1293–302. doi: 10.1002/mrm.22916
9. Mankodi A, Bishop CA, Auh S, Newbould RD, Fischbeck KH, Janiczek RL. Quantifying disease activity in fatty-infiltrated skeletal muscle by IDEAL-CPMG in Duchenne muscular dystrophy. *Neuromuscul Disord.* (2016) 26:650–8. doi: 10.1016/j.nmd.2016.07.013
10. Schlaeger S, Weidlich D, Klupp E, Montagnese F, Deschauer M, Schoser B, et al. Water T2 mapping in fatty infiltrated thigh muscles of patients with neuromuscular diseases using a T2-prepared 3D turbo spin echo with SPAIR. *J Magn Reson Imaging.* (2020) 51:1727–36. doi: 10.1002/jmri.27032
11. Sousa PL de, Vignaud A, Araújo EC de A, Carlier PG. Factors controlling T2 mapping from partially spoiled SSFP sequence: optimization for skeletal muscle characterization. *Magn Reson Med.* (2012) 67:1379–90. doi: 10.1002/mrm.23131
12. Klupp E, Weidlich D, Schlaeger S, Baum T, Cervantes B, Deschauer M, et al. B1-insensitive T2 mapping of healthy thigh muscles using a T2-prepared 3D TSE sequence. *PLoS ONE.* (2017) 12:e0171337. doi: 10.1371/journal.pone.0171337

13. Koolstra K, Webb AG, Veeger TTJ, Kan HE, Koken P, Börnert P. Water–fat separation in spiral magnetic resonance fingerprinting for high temporal resolution tissue relaxation time quantification in muscle. *Magn Reson Med.* (2020) 84:646–62. doi: 10.1002/mrm.28143
14. Marty B, Baudin P-Y, Reyngoudt H, Azzabou N, Araujo ECA, Carlier PG, et al. Simultaneous muscle water T2 and fat fraction mapping using transverse relaxometry with stimulated echo compensation. *NMR Biomed.* (2016) 29:431–43. doi: 10.1002/nbm.3459
15. Keene KR, Beenakker J-WM, Hooijmans MT, Naarding KJ, Niks EH, Otto LAM, et al. T2 relaxation-time mapping in healthy and diseased skeletal muscle using extended phase graph algorithms. *Magn Reson Med.* (2020) 84:2656–70. doi: 10.1002/mrm.28290
16. Strijkers GJ, Araujo ECA, Azzabou N, Bendahan D, Blamire A, Burakiewicz J, et al. Exploration of new contrasts, targets, and MR imaging and spectroscopy techniques for neuromuscular disease – A Workshop Report of Working Group 3 of the Biomedicine and Molecular Biosciences COST Action BM1304 MYO-MRI. *J Neuromuscul Dis.* (2019) 6:1–30. doi: 10.3233/JND-180333
17. Kan HE, Scheenen TWJ, Wohlgemuth M, Klomp DWJ, van Loosbroek-Wagenmans I, Padberg GW, et al. Quantitative MR imaging of individual muscle involvement in facioscapulohumeral muscular dystrophy. *Neuromuscul Disord.* (2009) 19:357–62. doi: 10.1016/j.nmd.2009.02.009
18. Schlaffke L, Rehmann R, Rohm M, Otto LAM, de Luca A, Burakiewicz J, et al. Multi-center evaluation of stability and reproducibility of quantitative MRI measures in healthy calf muscles. *NMR Biomed.* (2019) 32:e4119. doi: 10.1002/nbm.4119
19. Hooijmans MT, Baligand C, Froeling M, Verschuuren JJGM, Webb AG, Niks EH, et al. Multi-parametric MR shows increased T2 heterogeneity in fat infiltrated muscles in Becker Muscular Dystrophy. In: *Proc Intl Soc Mag Reson Med 26.* Paris (2018). Available from: http://indexsmart.mirasmart.com/ISMRM2018/PDFfiles/0816.html (accessed January 15, 2021).
20. Otto LAM, Pol W-L van der, Schlaffke L, Wijngaarde CA, Stam M, Wadman RI, et al. Quantitative MRI of skeletal muscle in a cross-sectional cohort of patients with spinal muscular atrophy types 2 and 3. *NMR Biomed.* (2020) 33:e4357. doi: 10.1002/nbm.4357
21. Heskamp L, Okkersen K, van Nimwegen M, Ploegmakers MJ, Bassez G, Deux J-F, et al. Quantitative muscle MRI depicts increased muscle mass after a behavioral change in myotonic dystrophy type 1. *Radiology.* (2020) 297:132–42. doi: 10.1148/radiol.2020192518
22. Keene KR, Vught L, Velde NM, Ciggaar IA, Notting IC, Genders SW, et al. The feasibility of quantitative MRI of extra-ocular muscles in myasthenia gravis and Graves' orbitopathy. *NMR Biomed.* (2021) 34:e4407. doi: 10.1002/nbm.4407
23. Carlier PG, Marty B, Scheidegger O, Loureiro de Sousa P, Baudin P-Y, Snezhko E, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging and spectroscopy as an outcome measure for clinical trials. *J Neuromuscul Dis.* (2016) 3:1–28. doi: 10.3233/JND-160145
24. Radunsky D, Blumenfeld-Katzir T, Volovyk O, Tal A, Barazany D, Tsarfaty G, et al. Analysis of magnetization transfer (MT) influence on quantitative mapping of T2 relaxation time. *Magn Reson Med.* (2019) 82:145–58. doi: 10.1002/mrm.27704
25. Weigel M, Helms G, Hennig J. Investigation and modeling of magnetization transfer effects in two-dimensional multislice turbo spin echo sequences with low constant or variable flip angles at 3 T. *Magn Reson Med.* (2010) 63:230–4. doi: 10.1002/mrm.22145
26. Melki PS, Mulkern RV. Magnetization transfer effects in multislice RARE sequences. *Magn Reson Med.* (1992) 24:189–95. doi: 10.1002/mrm.1910240122
27. Welch EB, Smith DS, Avison MJ, Berglund J, Kullberg J, Ahlström H. *Fattyriot - Final Winning Entry Of The 2012 Ismrm Challenge On Water-Fat Reconstruction*. Zenodo (2015). Available from: https://zenodo.org/record/16741 (accessed October 9, 2020).
28. Huang C, Altbach MI, El Fakhri G. Pattern recognition for rapid T2 mapping with stimulated echo compensation. *Magn Reson Imaging.* (2014) 32:969–74. doi: 10.1016/j.mri.2014.04.014
29. Weigel M. Extended phase graphs: dephasing, RF pulses, and echoes - pure and simple. *J Magn Reson Imaging.* (2015) 41:266–95. doi: 10.1002/jmri.24619
30. Hennig J. Echoes—how to generate, recognize, use or avoid them in MR-imaging sequences. Part I: fundamental and not so fundamental properties of spin echoes. *Concepts Magn Reson.* (1991) 3:125–43. doi: 10.1002/cmr.1820030302
31. Carr HY, Purcell EM. Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys Rev.* (1954) 94:630–8. doi: 10.1103/PhysRev.94.630
32. Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. *Rev Sci Instrum.* (1958) 29:688–91. doi: 10.1063/1.1716296
33. Pauly J, Roux PL, Nishimura D, Macovski A. Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm (NMR imaging). *IEEE Trans Med Imaging.* (1991) 10:53–65. doi: 10.1109/42.75611
34. Tasca G, Monforte M, Ottaviani P, Pelliccioni M, Frusciante R, Laschena F, et al. Magnetic resonance imaging in a large cohort of facioscapulohumeral muscular dystrophy patients: pattern refinement and implications for clinical trials. *Ann Neurol.* (2016) 79:854–64. doi: 10.1002/ana.24640
35. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing (2008). Available from: http://www.R-project.org (accessed February 15, 2021).
36. Froeling M. QMRTools: a mathematica toolbox for quantitative MRI analysis. *J Open Source Softw.* (2019) 4:1204. doi: 10.21105/joss.01204
37. Roujol S, Weingärtner S, Foppa M, Chow K, Kawaji K, Ngo LH, et al. Accuracy, precision, and reproducibility of four T1 mapping sequences: a head-to-head comparison of MOLLI. ShMOLLI, SASHA, and SAPPHIRE. *Radiology.* (2014) 272:683–9. doi: 10.1148/radiol.14140296
38. Baeßler B, Schaarschmidt F, Stehning C, Schnackenburg B, Giolda A, Maintz D, et al. Reproducibility of three different cardiac T2-mapping sequences at 1.5T. *J Magn Reson Imaging.* (2016) 44:1168–78. doi: 10.1002/jmri.25258
39. Schlaeger S, Weidlich D, Klupp E, Montagnese F, Deschauer M, Schoser B, et al. Decreased water T2 in fatty infiltrated skeletal muscles of patients with neuromuscular diseases. *NMR Biomed.* (2019) 32:e4111. doi: 10.1002/nbm.4111
40. Hardy PA, Henkelman RM, Bishop JE, Poon ECS, Plewes DB. Why fat is bright in rare and fast spin-echo imaging. *J Magn Reson Imaging.* (1992) 2:533–40. doi: 10.1002/jmri.1880020511

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 01 March 2021
doi: 10.3389/fneur.2021.621257

Correlation Between Respiratory Accessory Muscles and Diaphragm Pillars MRI and Pulmonary Function Test in Late-Onset Pompe Disease Patients

*David Reyes-Leiva[1,2], Jorge Alonso-Pérez[1,2], Mercedes Mayos[3], Claudia Nuñez-Peralta[4], Jaume Llauger[4], Izaskun Belmonte[5], Irene Pedrosa-Hernández[5], Sonia Segovia[1,2] and Jordi Díaz-Manera[1,2,6]**

[1] Neuromuscular Disorders Unit, Neurology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [2] Centro de Investigación Biomédica en Red en Enfermedades Raras, Madrid, Spain, [3] Pneumology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [4] Radiology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [5] Rehabilitation and Physiotherapy Department, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain, [6] John Walton Muscular Dystrophy Research Center, Newcastle University, Newcastle, United Kingdom

OPEN ACCESS

Edited by:
Edoardo Malfatti,
INSERM U1179 Handicap neuromusculaire: Physiopathologie, Biothérapie et Pharmacologie appliquées (END-ICAP), France

Reviewed by:
Corrado Italo Angelini,
University of Padua, Italy
Antonio Toscano,
University of Messina, Italy

***Correspondence:**
Jordi Díaz-Manera
jordi.diaz-manera@newcastle.ac.uk

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: 25 October 2020
Accepted: 13 January 2021
Published: 01 March 2021

Citation:
Reyes-Leiva D, Alonso-Pérez J, Mayos M, Nuñez-Peralta C, Llauger J, Belmonte I, Pedrosa-Hernández I, Segovia S and Díaz-Manera J (2021) Correlation Between Respiratory Accessory Muscles and Diaphragm Pillars MRI and Pulmonary Function Test in Late-Onset Pompe Disease Patients. Front. Neurol. 12:621257. doi: 10.3389/fneur.2021.621257

Objectives: Pompe disease is a rare genetic disease produced by mutations in the GAA gene leading to progressive skeletal and respiratory muscle weakness. T1-weighted magnetic resonance imaging is useful to identify fatty replacement in skeletal muscles of late-onset Pompe disease (LOPD) patients. Previous studies have shown that replacement by fat correlates with worse results of muscle function tests. Our aim was to investigate if fat replacement of muscles involved in the ventilation process correlated with results of the spirometry and predicted respiratory muscle impairment in LOPD patients over time.

Materials and Methods: We studied a cohort of 36 LOPD patients followed up annually in our center for a period of 4 years. We quantified muscle fat replacement using Mercuri score of the thoracic paraspinal and abdominal muscles and the pillars of the diaphragm. We correlated the combined Mercuri scores of these areas with spirometry results and the need of respiratory support.

Results: We found a statistically significant correlation (Spearman test, $p < 0.05$; coefficient of correlation > 0.6) between forced vital capacity seated and lying and fat fraction score of all muscle groups studied. The group of patients who needed respiratory support had higher fat fraction scores than patients not requiring ventilatory support. Higher fat replacement in these areas correlated with worse progression in spirometry values over time.

Conclusions: Fat replacement of paraspinal, abdominal, and trunk muscles correlates with results of spirometry and is able to predict worsening in respiratory muscle function

tests that could lead to an emerging ventilatory dysfunction. Therefore, the identification of fat replacement in these muscle groups should lead to a closer monitorization of patients. Radiologic evaluation of diaphragm pillars in T1-weighted imaging axial sequences could also be helpful to predict respiratory insufficiency.

Keywords: Pompe disease, glycogen storage disease type II, MRI, muscular MRI, respiratory insufficiency, accessory respiratory muscles, diaphragm pillars

INTRODUCTION

Pompe disease or glycogen storage disease type II is a rare genetic disease produced by mutations in the *GAA* gene encoding the enzyme acid alpha glucosidase (AAG). This enzyme catalyzes glycogen into glucose inside the lysosomes. Mutations in the *GAA* gene lead to accumulation of lysosomes loaded with glycogen and autophagic vacuoles inside cells in several tissues but especially in skeletal, smooth, and heart muscles (1). This accumulation triggers only partially known intracellular molecular pathways leading to necrosis and substitution of muscle fibers by fat and fibrous tissue that eventually produces permanent skeletal and respiratory muscle weakness.

Pompe disease is differentiated in two main phenotypes. In infantile-onset Pompe disease, symptoms start during the first months of life and are characterized by hypotonia, general muscle weakness, respiratory impairment, and hypertrophic cardiomyopathy. This phenotype is very severe with an ominous prognosis and is associated with a very low or absent expression of AAG (2). In contrast, in late-onset Pompe disease (LOPD), symptoms start in any moment after 1 year old, being characteristic a slowly progressive weakness affecting proximal muscles of the limbs and/or axial muscles associated or not with respiratory muscle involvement and elevated creatine kinase (CK) values in blood tests in most patients (2, 3). The clinical spectrum in LOPD is heterogeneous, ranging from isolated asymptomatic hyperCKemia to severe muscle weakness and respiratory insufficiency requiring invasive ventilatory support. Enzymatic replacement therapy (ERT) with alglucosidase alfa (Myozyme/Lumizyme®, Sanofi-Genzyme, Cambridge, MA, USA) is the only approved treatment for this disease (4). A single placebo-controlled clinical trials and several open-label studies have already shown the efficacy of ERT maintaining muscle and respiratory function over time (5–8).

Respiratory impairment is a common manifestation of LOPD. Approximately 60–80% of LOPD patients develop respiratory insufficiency during their lifetime (9). Ventilation is a complex process that involves several muscles of the trunk, pharynx, and larynx to accomplish an adequate inspiration and expiration cycle. Diaphragm and intercostal muscles are considered primary respiratory muscles playing a key role expanding the thoracic cavity during inspiration. Other muscles such as scalenus, serratus, sternocleidomastoid, and pectoralis support inspiration when diaphragm and intercostals are weak enough to not achieve an appropriate inspiration (10, 11). Respiratory insufficiency in LOPD patients is mainly produced by diaphragm dysfunction that should be checked routinely with pulmonary function test (PFT) using spirometry. A difference of 10% or more in forced vital capacity (FVC) in sitting position compared to lying is considered a sign of diaphragm muscle weakness (12). Other parameters assessed with conventional spirometry include maximal inspiratory pressure (MIP) that also correlates with diaphragm performance. However, PFTs are often influenced by patient's motivation, and to obtain a perfect test is not always easy from a technical point of view.

In recent years, muscle magnetic resonance imaging (MRI) using T1-weighted imaging (T1w) has proved to be useful identifying skeletal muscle fat replacement in several neuromuscular diseases, including LOPD. Previous studies have shown that trunk and lower limbs muscles are commonly replaced by fat in LOPD patients and that the amount of fat present in these muscles correlates with the degree of muscle impairment measured using different muscle function tests (13, 14). Therefore, muscle MRI could be considered a good biomarker of muscle function in this disease useful to distinguish between mild and severe patients (15). Based on these previous results, we wondered if fat replacement of the muscles involved in the ventilation process correlated with spirometry values or predicted respiratory muscle impairment in LOPD patients over time.

MATERIALS AND METHODS

Description of the Study Population and PFT

This study was a prospective longitudinal cohort study involving 36 genetically confirmed LOPD patients performed at Hospital de la Santa Creu i Sant Pau (HSCP) in Barcelona from December 2013 to June 2018. The clinical and genetic features of this cohort have been previously described (15). All patients were seen once per year, and the assessment included a clinical interview, several muscle function tests, a conventional spirometry, and a whole-body muscle MRI using T1w sequence. We obtained PFT during the 4 years' duration of the study, and MRI T1w whole body during the first 3 years. On the fourth year, we changed the MRI protocol, and only images of the lower limbs were obtained, so we did not use those images in the present study. The study was approved by the HSCP Ethics Committee, and all participants signed a consent form. The study was registered in ClinicalTrials.gov with the identifier NCT01914536.

Inclusion criteria for the study were (1) a diagnosis of LOPD according to the recommendations made by the European Pompe Consortium: reduced enzymatic activity on fibroblasts, blood leukocytes or skeletal muscle and/or the presence of two known mutations in the GAA gene (16); (2) no contraindications

for the MRI; and (3) willingness to complete muscle function test and respiratory assessment. Asymptomatic patients who only had isolated hyperCKemia without clear skeletal muscle weakness in clinical examination or respiratory symptoms were also included. The criteria requested by health authorities to approve ERT in Spain vary from one region to another, but are commonly based on the presence of skeletal and/or respiratory muscular weakness in clinical examination. Isolated hyperCKemia associated or not to fatigue or muscle pain is not considered a criterion to start ERT.

Muscle function tests and spirometry were performed by two physiotherapists with a long experience in the assessment of neuromuscular patients (I.B., I.P.-H.). In terms of spirometry, we measured FVC seated and lying and MIP and maximal expiratory

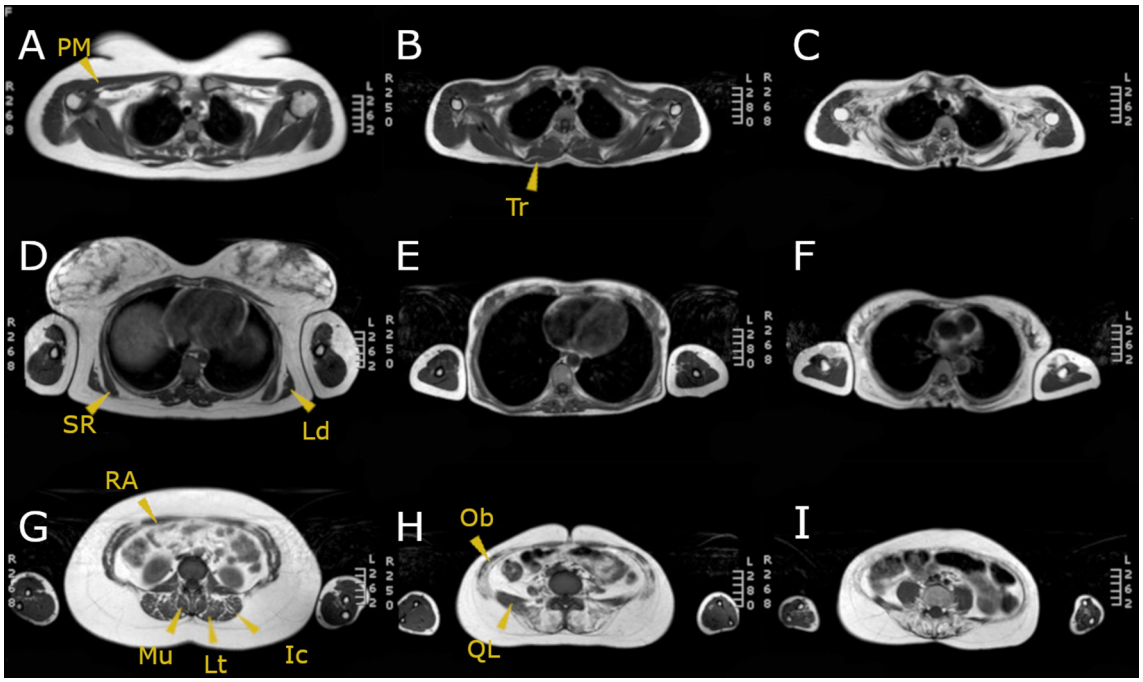


FIGURE 1 | MRI T1w of three LOPD patients: **(A,D,G)** correspond to a patient with mild involvement; **(B,E,H)** correspond to a patient with moderate muscular involvement; **(C,F,I)** correspond to a patient with severe muscular involvement. Yellow arrowheads point the muscles assessed. PM, pectoralis major; SR, serratus; Ld, latissimus dorsi; RA, rectus abdominis; Mu, multifidus; Lt, longissimus thoracis; Ic, iliocostalis; Tr, trapezius; Ob, obliquus externus and internus; QL, quadratus lumborum.

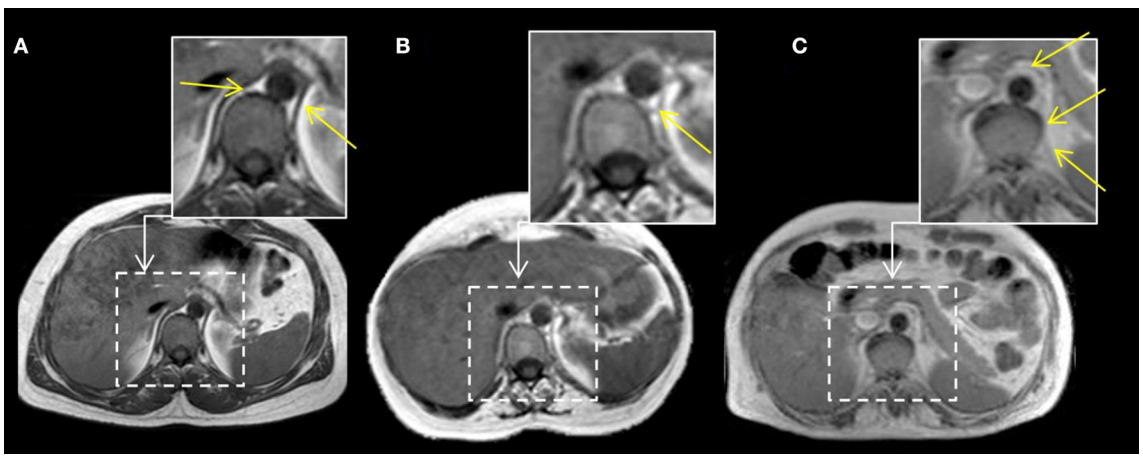


FIGURE 2 | MRI T1w abdominal sections of three LOPD patients. Yellow arrows point to both diaphragm pillars. In case **(A)**, right and left diaphragm pillars are visible and do not show fatty replacement (Mercuri score 0). In case **(B)**, left diaphragm pillar is visible and partially infiltrated, whereas right diaphragm pillar is absent (left pillar Mercuri score 3 points). In case **(C)**, we can observe fully replaced diaphragm pillars at both sides (Mercuri score 4 points).

pressure (MEP) with the Carefusion Microlab ML 3500 MK8 spirometer (Carefusion, Yorba Linda, CA, USA).

Muscle Imaging

The 36 LOPD patients were imaged in a 1.5-T field magnetic resonance system (1.5-T Achieva, Philips, Eindhoven, the Netherlands) at HSCSP in Barcelona during the whole duration of the study. Patients were imaged in supine position with the legs stretched out. Whole-body axial T1w images were obtained. The images were analyzed by three members of our team (D. R.-L., C. N.-P., J. D.-M.) who quantified fatty muscle replacement using the modified version of the Mercuri score published by Dr. Fischer (17): no fatty replacement: 0 point; mild fatty replacement or traces of infiltration on T1w scores, 1 point; fatty replacement in <50% of the muscle scores, 2 points; fatty replacement in more than 50% of the muscle scores, 3 points; and end-stage appearance with the whole muscle replaced by fatty tissue scores, 4 points.

We analyzed fat replacement of the so-called respiratory accessory muscles both in the right and left sides of the body as shown in **Figure 1**. We divided respiratory accessory

TABLE 1 | Clinical features of the cohort.

Patient	Sex	Age at diagnosis	Age at baseline visit	Age at ERT	Clinical phenotype	Age at ventilation	Mechanical ventilation at Baseline Visit	Mechanical ventilation at last visit
1	Female	38	50	47	MW	–	No	No
2	Female	20	48	39	MW + RS	34	Yes (noninvasive)	Yes (noninvasive)
3	Female	–	26	No	HyperCKemia	–	No	No
4	Female	56	63	59	MW	–	No	No
5	Female	20	45	42	MW	–	No	No
6	Female	36	51	47	MW	–	No	No
7	Male	62	66	67	MW + RS	67	No	Yes (noninvasive)
8	Female	49	59	52	MW	–	No	No
9	Female	29	55	48	MW	–	No	No
10	Male	13	42	39	MW + RS	38	Yes (noninvasive)	Yes (noninvasive)
11	Female	23	31	24	MW + RS	23	Yes (noninvasive)	Yes (noninvasive)
12	Female	20	46	39	MW	–	No	No
13	Male	43	47	45	MW + RS	44	Yes (noninvasive)	Yes (noninvasive)
14	Male	41	51	45	MW + RS	50	Yes (noninvasive)	Yes (noninvasive)
15	Female	35	51	46	MW + RS	46	Yes (noninvasive)	Yes (noninvasive)
16	Male	–	22	No	HyperCKemia	–	No	No
17	Male	–	14	No	HyperCKemia	–	No	No
18	Female	40	65	64	MW	–	No	No
19	Female	24	35	29	MW	–	No	No
20	Female	29	40	40	MW + RS	39	Yes (noninvasive)	Yes (noninvasive)
21	Female	20	52	45	MW	–	No	No
22	Male	47	64	57	MW + RS	56	Yes (noninvasive)	Yes (noninvasive)
23	Male	2	8	No	HyperCKemia	–	No	No
24	Female	42	57	55	MW + RS	54	Yes (noninvasive)	Yes (noninvasive)
25	Male	41	46	43	MW + RS	43	Yes (noninvasive)	Yes (noninvasive)
26	Male	35	51	51	MW + RS	49	Yes (noninvasive)	Yes (noninvasive)
27	Male	35	51	No	MW	–	No	No
28	Male	20	43	43	MW	–	No	No
29	Female	40	54	48	MW	–	No	No
30	Male	–	12	No	HyperCKemia	–	No	No
31	Male	49	51	No	HyperCKemia + OA.	–	No	No
32	Male	39	43	No	MW	–	No	No
33	Female	16	20	No	HyperCKemia	–	No	No
34	Female	28	39	No	MW	–	No	No
35	Female	35	35	No	HyperCKemia	–	No	No
36	Male	44	49	No	HyperCKemia	–	No	No

Age is expressed in years. HyperCKemia patients did not have other clinical manifestations except of patient 31. ERT, enzyme replacement therapy; MW, muscular weakness; RS, respiratory symptoms; OA, obstructive apnea.

TABLE 2 | Mean values of the compound MRI scores at every visit.

	Visit 1, mean (SD)	Visit 2, mean (SD)	Visit 3, mean (SD)
Abdominal score	19.52 (12.5)	19.77 (12.5)	21.5 (11.86)
Paraspinal score	19.11 (11.17)	20.08 (11.4)	21.78 (10.77)
Thorax score	9.08 (9.57)	9.97 (9.53)	11.55 (10.16)
Global score	48.32 (29.78)	49.83 (30.58)	54.89 (29.60)

SD, standard deviation.

muscles in three groups: thoracic muscles, paraspinal muscles, and abdominal muscles. We calculated a compound modified Mercuri score of these areas adding the individual Mercuri score for each of those muscles from both sides (left and right) and eventually obtaining four compound scores:

- Thoracic muscle score: Serratus L (0–4) + serratus R (0–4) + latissimus dorsi L (0–4) + latissimus dorsi R (0–4) + pectoralis major L (0–4) + pectoralis major R (0–4) + scalenus L (0–4) + scalenus R (0–4) + trapezius L (0–4) + trapezius R (0–4). This score ranged from 0 to 40.
- Paraspinal muscle score: Multifidus L (0–4) + multifidus R (0–4) + longissimus thoracis L (0–4) + longissimus thoracis R (0–4) + quadratus lumborum L (0–4) + quadratus lumborum R (0–4) + iliocostalis L (0–4) + iliocostalis R (0–4). This score ranged from 0 to 32
- Abdominal muscle score: Obliquus externus L (0–4)+ obli quus externus R (0–4) + obliquus internus L (0–4) + obliquus internus R (0–4) + transversus abdominis L (0–4) + transversus abdominis R (0–4)+ rectus abdominis L (0–4) + rectus abdominis R (0–4). This score ranged from 0 to 32.
- Global Score: Thoracic muscle score (0–40) + paraspinal muscle score (0–32) + Abdominal muscle score (0–32). This last score ranged from 0 to 104.

Furthermore, we assessed the involvement of the diaphragm by analyzing fatty replacement of the diaphragm pillars, which are two musculotendinous structures easily identifiable surrounding the anterior lumbar spine as shown in **Figure 2**.

Statistics

We confirmed the scores obtained were normally distributed using Kolmogorov–Smirnov test, and consequently, we used parametric statistic studies to analyze the data. We used Student *t*-test to study if the differences observed in the fatty replacement scores obtained between ventilated and nonventilated patients were significant. To analyze differences observed over time in fatty replacement, we used a mixed linear model and applied Greenhouse–Geisser test. We used Pearson test to correlate respiratory measurements with the degree of muscle involvement, and we considered that the correlation was good if the correlation coefficient were higher than 0.6. Hierarchical analysis and graphical representation as a heatmap was performed using R software, V.3.1.3. Significance was set up at *p*-values lower than 0.05. Statistical studies were performed with SPSS for MAC computers (version 21, SPSS Inc., Chicago, IL).

TABLE 3 | Mean values of the MRI scores and Mercuri score of diaphragm pillars classified in ventilated and nonventilated patients.

	Noninvasive mechanical ventilation = 12 (SD)	Nonventilated patients = 24 (SD)	*p*-value
Abdominal score	26.16 (6.8)	16.95 (13.62)	0.013
Paraspinal score	26.8 (3.68)	15.54 (11.94)	0.001
Thorax score	12.7 (9.16)	7.7 (9.64)	0.155
Global score	65.75 (14.32)	40.6 (32.53)	0.006
Diaphragm pillar left	2.66 (1.7)	1.2 (1.6)	0.036
Diaphragm pillar right	2.66 (1.7)	1.2 (1.6)	0.036

The last column shows results of Student t-test (significant $p < 0.05$). SD = standard deviation.

RESULTS

Cohort Description

Clinical data of the 36 LOPD patients enrolled in the study have been previously reported and are displayed in **Table 1** (15). Twenty patients were female (55.5%). Mean age of the patients at the start of the study was 43.9 ± 14.8 years. At the baseline visit, 23 patients were on ERT with Myozyme. One patient (patient 7) started ERT during the follow-up. Mean duration of ERT treatment at baseline was 4.30 ± 2.69 years. Eight patients were considered asymptomatic as they only presented hyperCKemia and were not receiving ERT. The most common clinical feature of the cohort was muscular weakness involving the lower limbs. Eleven patients needed noninvasive nocturnal mechanical ventilation at the start of our study. Patient 7 started noninvasive nocturnal mechanical ventilation during the study.

MRI T1w Analysis

Both abdominal and paraspinal muscles were more commonly and severely affected than thoracic muscles in our cohort as shown in **Tables 2**, **3** and in **Figure 3**. The muscles more commonly replaced by fat were multifidus, longissimus, iliocostalis, rectus abdominis, the oblique muscles and transversus abdominis. Latissimus dorsi was the thoracic muscle more common and severely affected, while there was a clear heterogeneity in the involvement of the remaining thoracic muscles among patients as shown in **Figure 3**. We observed significant differences in the degree of muscle fat replacement between ventilated and nonventilated patients in abdominal and paraspinal muscles but not in thoracic muscles at baseline (**Table 3**). It is noteworthy that the pillars of the diaphragm were more severely affected in ventilated than in nonventilated patients (**Table 3**) as shown in **Figure 2**. There was a significant progressive increase in the degree of fat replacement of thoracic, abdominal, and paraspinal muscles during the follow-up (Greenhouse Geisser, $p < 0.001$). Interestingly, some of the presymptomatic patients have mild signs of fatty replacement especially in paraspinal and abdominal muscles. It is noteworthy that one of the considered presymptomatic patients had more consistent changes affecting also thoracic muscles. However, this patient did not complain of any muscular symptom.

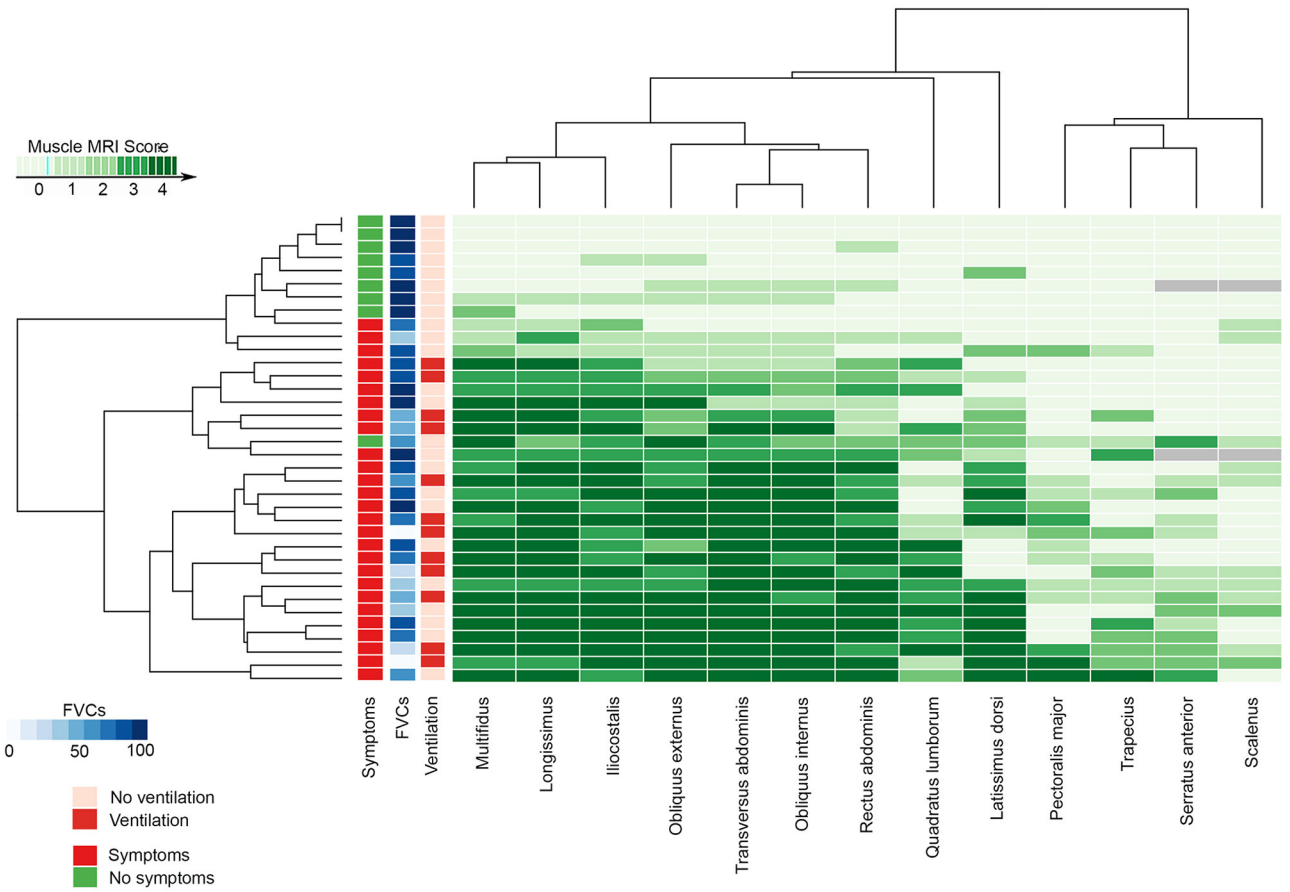


FIGURE 3 | The heatmap shows the MRI Mercuri score value for all the muscles studied, the presence of muscular weakness (symptomatic) or hyperCKemia alone (no symptomatic), its relationship with the FVC at base time, and the needs of noninvasive mechanical ventilation in every patient. FVC, forced vital capacity; MRI, magnetic resonance imaging.

Pulmonary Function Test Analysis

There were significant differences in FVC percentage predicted measured seated or lying between ventilated and nonventilated patients (Student *t*-test, $p = 0.001$ for both measurements). Moreover, we identified a significant difference in FVC seated between baseline and last visit both in ventilated and nonventilated patients (**Table 4**). We did not observe significant differences in FVC in presymptomatic patients between baseline and last visit. The MIP and MEP values at baseline visit were also studied obtaining a mean MIP of 70.52 ± 28.95% and a mean MEP 72.98 ± 31.77% in all the patients.

Correlation Test

We observed significant correlation between global, thoracic, abdominal, and paraspinal scores and FVC percentage predicted seated and lying at baseline. Additionally, there were significant correlations between fat replacement of the diaphragm pillars and FVC percentage predicted seated and lying and MIP (**Table 5**).

We also observed a significant correlation between baseline MRI score and changes in FVC percentage predicted after 3 years of follow-up for global ($p = 0.013$, $r = -0.506$), abdominal $p = 0.017$, $r = -0.602$), and paraspinal ($p = 0.013$, $r = -0.542$) MRI Score. Although significant, the correlations with thoracic MRI score and fat replacement of the diaphragm pillar were low: $r = -0.233$ and $r = -0.437$, respectively.

It is noteworthy that one of the patients ventilated (patient 24), with a clear diaphragmatic dysfunction characterized by a drop in FVC when lying higher than 20%, had mild abdominal but severe paraspinal fat replacement (abdominal score: 10, paraspinal score: 28) and normal diaphragm pillars. Additionally, we detected patients (patients 6, 9, 14, 29, and 31) who had mild to moderate abdominal fat replacement with a severe paraspinal fat replacement but with diaphragm pillars nonaffected and who did not show diaphragm dysfunction (FVC Seated - FVC lying <10%).

We want to highlight the individual case of patient 14 who is a 51-year-old male patient who was already using nocturnal ventilation when recruited in this study. The patient was diagnosed as having sleep apnea 3 years before being recruited and was on nocturnal ventilation 8 h per night at the beginning of the study, and this remained unmodified the whole duration of

the study. The predicted FVC seated was 78%, and the predicted FVC lying was 83% at baseline. At last visit, predicted FVC seated was 69%, and FVC lying was 64%. These differences were not compatible with diaphragmatic dysfunction using the definition mentioned above (a drop in FVC higher than 10% between seating and lying), but a worsening tendency was noticed. In this patient, muscle MRI showed a severe fatty replacement of paraspinal and abdominal muscles that was not modified throughout the whole duration of the study. We have included an MRI of this patient at first visit in **Figure 4**. The MRI scores measured for this patient on the baseline visit were 30 points in abdominal score, 25 points in paraspinal score, and 18 points in thoracic score, so the Global score at the first visit was of 73 points. In the last visit, we identified a very mild worsening in paraspinal score (26 points) and in thoracic score (19 points) remaining unchanged the abdominal score. The global score was of 75 then.

DISCUSSION

In this study, we show that fat replacement of thoracic, abdominal, and paraspinal muscles correlates with low values of FVC percentage predicted both seated and lying in a large cohort of LOPD patients. Additionally, fat replacement of abdominal and paraspinal muscles predicted changes in FVC during the follow-up in our cohort. These results suggest that identification of fat replacement in these areas on the MRI should lead to a closer monitorization of respiratory muscle function over time.

It is well known that LOPD patients have a selective involvement of paraspinal, abdominal and proximal muscles of the lower limbs in T1w MRI sequences (14, 15). Here we have confirmed these previous results, but we have also found a good correlation between fat replacement and results of respiratory muscle function tests. It has been previously shown that the degree of fat replacement in a muscle correlates with results of specific function tests including assessment of muscle strength using handheld dynamometry or even assessment of muscle performance using timed tests such as the 6-min walking test (18–20). Accordingly, it is not surprising that fat replacement of the diaphragm or of the accessory respiratory muscles, such as thoracic muscles, has a good correlation with FVC. However, we have seen here that abdominal and paraspinal involvement is also associated with FVC, and it predicts changes in this measure over time. This also suggests that clinical examination findings such as a distended abdominal wall, Beevor sign (cephalic umbilical shift with abdominal muscular contraction due to abdominal weakness), or axial weakness, which are a common finding in LOPD, should also point to a closer monitorization of respiratory function (21).

In concordance with these results, we have observed significant differences in the degree of fat replacement between ventilated and nonventilated LOPD patients, suggesting that fat replacement of these muscles leads to a decrease in respiratory muscle function and eventually to the need of nocturnal respiratory support. All patients using noninvasive ventilation had some degree of fatty replacement in diaphragm pillars, except patient 14, who despite being ventilated did not meet the standard criteria of diaphragm dysfunction. Although diaphragm pillar grading is not a standardized procedure, and there is a wide variability between healthy volunteers and patients, we consider it is a valid measure linked to clinical involvement of the diaphragm.

We have identified significant differences in FVC between baseline and last visit in symptomatic patients treated despite being treated with ERT. Several other studies have also analyzed the progression of FVC over time in LODP patients treated with ERT, and the results are variable. Most of the studies coincide in a positive effect of starting ERT on FVC, as this value remains stable in most of the patients during the first 5 years of treatment (5, 8, 22). However, a recent study including patients followed

TABLE 4 | Mean values of the FVC measured seated and mean values of the differences between FVC seated and lying at the first and last visit.

	First visit, mean FVC% seated (SD)	Last visit, mean FVC% seated (SD)	*p*-value
All patients (36)	79.98% (21.43)	73.42 % (25.34)	0.002
Nonventilated patients (24)	88.23% (18.61)	80.94% (24.21)	0.017
Ventilated patients (12)	65.39% (12.91)	59.7% (14.06)	0.018
	First visit, mean FVC% S-L (SD)	**Last visit, mean FVC% S-L (SD)**	***p*-value**
All patients (36)	11.51 (12.27)	8.46 (7.42)	0.107
Nonventilated patients (24)	10.45 (11.48)	8.11 (6.71)	0.460
Ventilated patients (12)	23. 16 (15.00)	11.00 (8.93)	0.057

We compared the mean values from the first and last visit (p-values in the last column). FVC%, forced vital capacity percentage predicted; FVC% S-L, FVC% measured seated – FVC% measured lying; SD, standard deviation.

TABLE 5 | Correlations found between the compound scores calculated and the diaphragm pillar involvement and the FVC% seated and lying.

	FVC% seated	FVC% lying	MIP	MEP
Global score	$p = 0.002$ $R = -0.52$	$p = 0.004$ $R = -0.608$	$p = 0.772$	$p = 0.135$
Thorax score	$p = 0.001$ $R = -0.541$	$p = 0.003$ $R = -0.621$	$p = 0.372$	$p = 0.124$
Abdominal score	$p = 0.011$ $R = -0.523$	$p = 0.021$ $R = -0.394$	$p = 0.301$	$p = 0.248$
Paraspinal score	$p = 0.007$ $R = -0.456$	$p = 0.007$ $R = -0.570$	$p = 0.308$	$p = 0.309$
Diaphragm pillars	$p = 0.0001$ $R = -0.590$	$p = 0.001$ $R = -0.662$	$p = 0.006$ $R = -0.689$	$p = 0.59$

We are showing the results of the Pearson test: p-value and R value. FVC%, forced vital capacity percentage predicted; MIP, maximal inspiratory pressure; MEP, maximal expiratory pressure.

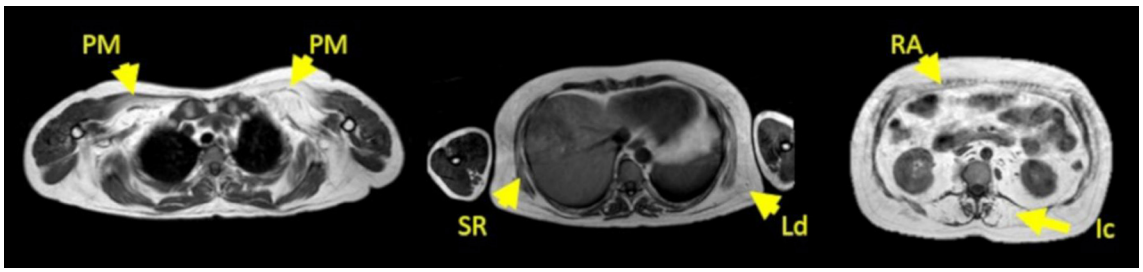


FIGURE 4 | MRI T1w of patient 14 at first visit. From left to right, first image corresponds to shoulder girdle level; second image corresponds to a thoracic level, and third image corresponds to an abdominal level. Yellow arrows point to the pectoralis major (PM), serratus (SR), latissimus dorsi (Ld), rectus abdominis (RA), and iliocostalis muscle (Ic). The MRI scores at this visit were 30 points in abdominal score, 25 points in paraspinal score, and 18 points in thoracic score. The Global score was of 73 points.

up for 10 years identified a mean decrease of more than 10% in FVC in patients treated, which is in agreement with our results, which are supported by an increase in fat replacement of respiratory muscles, suggesting that ERT is not able to completely stop progression of respiratory muscle degeneration in Pompe patients (7).

Other previous studies have identified a correlation between fatty replacement of abdominal muscles and diaphragm using computed tomography scan or MRI with spirometry results (23). We have confirmed these previous results, and additionally, we have also found a correlation with fatty replacement of paraspinal and thoracic muscles such as scalenus, trapezius, and pectoralis major. However, it is noteworthy that the correlation between FVC and fatty replacement of thoracic muscles was low, suggesting that they do not have a prominent role in respiratory function in LOPD patients, which is probably more dependent of the diaphragm and the abdominal muscles in these patients. Our protocol of MRI included only axial images, which unfortunately do not allow a good analysis of intercostalis muscles. We think that coronal images should be obtained if these muscles want to be analyzed. In recent years, several studies have also used MRI to study diaphragm function, and they have elegantly shown changes in morphology of thorax cavity and decrease in thorax dynamics during inspiration and expiration in LOPD patients (23–25). These results suggest a prominent involvement of the diaphragm, which is also supported by the frequent finding of a decrease in FVC percentage predicted higher than 10% when patients are lying compared to FVC seating and by the results of echography studies showing thinner diaphragms, which are replaced by fat (11, 26). The biggest limitation of our study is that we have used a semiquantitative scale to analyze fat replacement that is observer dependent and that have several limitations to identify consistent changes over time. At present, quantitative muscle MRI techniques, such as Dixon, are available and have demonstrated to identify better changes in fat replacement over time (27–29). However, it is important to take into account that Dixon studies of the trunk muscles are extremely challenging because of the movement artifacts of this area during normal respiration. To obtain images of good quality could be even more difficult in patients with respiratory involvement and orthopnea such as LOPD patients because the only way of avoiding motion artifacts using Dixon in this body areas is asking the patients to do repeated apneas that can last for 10–15 s (15, 30).

In conclusion, fat replacement of paraspinal, abdominal, and trunk muscles correlates with FVC percentage predicted seated and lying and is able to predict worsening in respiratory muscle function tests that could lead to an emerging ventilatory dysfunction. Therefore, the identification of fat replacement in these muscle groups should lead to a closer monitorization of patients.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics committee from Institut de Recerca Biomédica Hospital de la Santa Creu i Sant Pau. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DR-L, JA-P, CN-P, MM, JL, SS, and JD-M carried out the concept and design of the study. DR-L, JA-P, CN-P, JL, and JD-M performed the acquisition and analysis of data. JD-M performed the statistical analysis. DR-L, JA-P, IB, IP-H, and MM performed the acquisition and analysis of the PFT. DR-L, JA-P, CN-P, JL, SS, MM, and JD-M drafted of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was sponsored by a grant from Sanofi-Genzyme and with grants of the Fondo Europeo de Desarrollo Regional-FEDER, Instituto de Salud Carlos III (Spain) FIS (PI18/1525) to JD-M.

ACKNOWLEDGMENTS

We would like to thank all patients and relatives for the participation in this study. We also thank all members of the Spanish Pompe study group for their collaboration: Miguel Angel Barba-Romero MD (Hospital General de Albacete, Albacete, Spain); Joseba Barcena MD (Hospital Universitario Cruces, Baracaldo, Spain); María Rosario Carzorla MD (Hospital Puerta de Hierro, Majadahonda, Spain); Carlota Creus MD (Hospital Virgen de las Nieves, Granada, Spain); Jaume Coll-Cantí MD PhD (Hospital Germans Tries i Pujol, Badalona, Spain); Manuel Díaz MD (Hospital de Cabueñes, Gijón, Spain); Cristina Domínguez MD (Hospital 12 de Octubre, Madrid & Insituto de Investigación i+12, Madrid, Spain); Roberto Fernández Torrón MD (Hospital Universitario Donostia, Spain); María José García Antelo MD (Hospital Universitario A Coruña, A Coruña, Spain); JosepMaria Grau MD PhD (Hospital Clínic, Barcelona, Spain); Adolfo López deMunáin MD PhD (Hospital Universitario Donostia, Spain); Francisco Antonio Martínez-García MD (Hospital ClínicoUniversitario Virgen de la Arrixaca, Murcia, Spain); Yolanda Morgado, MD (Hospital Universitario Virgen de Valme, Sevilla, Spain); Antonio Moreno MD (Hospital Universitario Morales Meseguer, Murcia, Spain); Germán Morís, MD (Hospital Universitario de Asturias, Oviedo, Spain); Miguel Angel Muñoz-Blanco MD (Hospital GregorioMarañón, Madrid, Spain); Andres Nascimento MD (Hospital Sant Joan de Déu, Barcelona, Spain); Carmen Paradas PhD(Hospital Virgen del Rocío, Sevilla, Spain); José Luis Parajuá Pozo MD (Hospital de Can Mises, Ibiza, Spain); Montse Olivé Plana MD PhD (Hospital de Bellvitge, Barcelona, Spain); Arturo Robledo-Strauss MD (Hospital Juan Ramón Jiménez, Huelva, Spain); Íñigo Rojas-Marcos MD (Hospital Virgen del Rocio, Sevilla, Spain); Jose António Salazar MD (Hospital RegionalUniversitario deMálaga, Spain); Mercedes Usón MD (Hospital de Son Llátzer, Palma deMallorca, Spain).

REFERENCES

1. van der Ploeg A, Reuser A. Pompe's disease. *Lancet.* (2008) 372:1342–53. doi: 10.1016/S0140-6736(08)61555-X
2. Kohler L, Puertollano R, Raben N. Pompe disease: from basic science to therapy. *Neurotherapeutics.* (2018) 15:928–42. doi: 10.1007/s13311-018-0655-y
3. Bembi B, Cerini E, Danesino C, Donati M, Gasperini S, Morandi L, et al. Diagnosis of glycogenosis type II. *Neurology.* (2008) 71 (Suppl. 2):S4–S11. doi: 10.1212/WNL.0b013e31818da91e
4. van der Ploeg AT, Clemens PR, Corzo D, Escolar DM, Florence J, Groeneveld GJ, et al. A randomized study of alglucosidase alfa in late-onset pompe's disease. *N Engl J Med.* (2010) 362:1396–406. doi: 10.1056/NEJMoa0909859
5. Kuperus E, Kruijshaar M, Wens S, de Vries J, Favejee M, van der Meijden J, et al. Long-term benefit of enzyme replacement therapy in Pompe disease. *Neurology.* (2017) 89:2365–73. doi: 10.1212/WNL.0000000000004711
6. van der Meijden J, Kruijshaar M, Harlaar L, Rizopoulos D, van der Beek N, van der Ploeg A. Long-term follow-up of 17 patients with childhood Pompe disease treated with enzyme replacement therapy. *J Inherit Metab Dis.* (2018) 41:1205–14. doi: 10.1007/s10545-018-0166-3
7. Harlaar L, Hogrel J, Perniconi B, Kruijshaar M, Rizopoulos D, Taouagh N, et al. Large variation in effects during 10 years of enzyme therapy in adults with Pompe disease. *Neurology.* (2019) 93:e1756–67. doi: 10.1212/WNL.0000000000008441
8. Schoser B, Stewart A, Kanters S, Hamed A, Jansen J, Chan K, et al. Survival and long-term outcomes in late-onset Pompe disease following alglucosidase alfa treatment: a systematic review and meta-analysis. *J Neurol.* (2016) 264:621–30. doi: 10.1007/s00415-016-8219-8
9. Hagemans ML, Winkel LP, Van Doorn PA, Hop WJ, Loonen MC, Reuser AJ, et al. Clinical manifestation and natural course of late-onset Pompe's disease in 54 Dutch patients. *Brain*. (2005) 128:671–7. doi: 10.1093/brain/awh384
10. Wijdicks EFM. The neurology of acutely failing respiratory mechanics. *Ann Neurol.* (2017) 81:485–94. doi: 10.1002/ana.24908
11. Harlaar L, Ciet P, van der Ploeg AT, Brusse E, van der Beek NAME, Wielopolski PA, et al. Imaging of respiratory muscles in neuromuscular disease: a review. *Neuromuscul Disord.* (2018) 28:246–56. doi: 10.1016/j.nmd.2017.11.010
12. McCool F, Tzelepis G. Dysfunction of the diaphragm. *N Engl J Med.* (2012) 366:932–42. doi: 10.1056/NEJMra1007236
13. Pichiecchio A, Uggetti C, Ravaglia S, Egitto M, Rossi M, Sandrini G, et al. Muscle MRI in adult-onset acid maltase deficiency. *Neuromuscul Disord.* (2004) 14:51–5. doi: 10.1016/j.nmd.2003.08.003
14. Carlier R, Laforet P, Wary C, Mompoint D, Laloui K, Pellegrini N, et al. Whole-body muscle MRI in 20 patients suffering from late onset Pompe disease: involvement patterns. *Neuromuscul Disord.* (2011) 21:791–9. doi: 10.1016/j.nmd.2011.06.748
15. Figueroa-Bonaparte S, Segovia S, Llauger J, Belmonte I, Pedrosa I, Alejaldre A, et al. Muscle MRI findings in childhood/adult onset pompe disease correlate with muscle function. *PLoS ONE.* (2016) 11:e0163493. doi: 10.1371/journal.pone.0163493
16. van der Ploeg A, Kruijshaar M, Toscano A, Laforêt P, Angelini C, Lachmann R, et al. European consensus for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a 10-year experience. *Eur J Neurol.* (2017) 24:768–e31. doi: 10.1111/ene.13285
17. Fischer D, Kley R, Strach K, Meyer C, Sommer T, Eger K, et al. Distinct muscle imaging patterns in myofibrillar myopathies. *Neurology.* (2008) 71:758–65. doi: 10.1212/01.wnl.0000324927.28817.9b
18. Nuñez-Peralta C, Alonso-Pérez J, Llauger J, Segovia S, Montesinos P, Belmonte I, et al. Follow-up of late-onset Pompe disease patients with muscle magnetic resonance imaging reveals increase in fat replacement in skeletal muscles. *J Cachexia Sarcopenia Muscle.* (2020) 11:1032–46. doi: 10.1002/jcsm.12555
19. Alonso-Jimenez A, Kroon RHMJM, Alejaldre-Monforte A, Nuñez-Peralta C, Horlings CGC, van Engelen BGM, et al. Muscle MRI in a large cohort of patients with oculopharyngeal muscular dystrophy. *J Neurol Neurosurg Psychiatry.* (2019) 90:576–85. doi: 10.1136/jnnp-2018-319578
20. Diaz-Manera J, Fernandez-Torron R, LLauger J, James MK, Mayhew A, Smith FE, et al. Muscle MRI in patients with dysferlinopathy: pattern recognition and implications for clinical trials. *J Neurol Neurosurg Psychiatry.* (2018) 89:1071–81. doi: 10.1136/jnnp-2017-317488
21. Garibaldi M, Diaz-Manera J, Gallardo E, Antonini G. Teaching video neuroimages: the Beevor sign in late-onset Pompe disease. *Neurology.* (2016) 86:e250–1. doi: 10.1212/WNL.0000000000002772
22. Toscano A, Schoser B. Enzyme replacement therapy in late-onset Pompe disease: a systematic literature review. *J Neurol.* (2012) 260:951–9. doi: 10.1007/s00415-012-6636-x
23. Gaeta M, Barca E, Ruggeri P, Minutoli F, Rodolico C, Mazziotti S, et al. Late-onset Pompe disease (LOPD): correlations between respiratory muscles CT and MRI features and pulmonary function. *Mol Genet Metab.* (2013) 110:290–6. doi: 10.1016/j.ymgme.2013.06.023
24. Wens S, Ciet P, Perez-Rovira A, Logie K, Salamon E, Wielopolski P, et al. Lung MRI and impairment of diaphragmatic function in Pompe disease. *BMC Pulm Med.* (2015) 15:54. doi: 10.1186/s12890-015-0058-3
25. Gaeta M, Musumeci O, Mondello S, Ruggeri P, Montagnese F, Cucinotta M, et al. Clinical and pathophysiological clues of respiratory dysfunction in

late-onset Pompe disease: new insights from a comparative study by MRI and respiratory function assessment. *Neuromuscul Disord.* (2015) 25:852–8. doi: 10.1016/j.nmd.2015.09.003
26. Ruggeri P, Lo Monaco L, Musumeci O, Tavilla G, Gaeta M, Caramori G, et al. Ultrasound assessment of diaphragm function in patients with late-onset Pompe disease. *Neurol Sci.* (2020) 41:2175–84. doi: 10.1007/s10072-020-04316-6
27. Nuñez-Peralta C, Alonso-Pérez J, Díaz-Manera J. The increasing role of muscle MRI to monitor changes over time in untreated and treated muscle diseases. *Curr Opin Neurol.* (2020) 33:611–20. doi: 10.1097/WCO.0000000000000851
28. Dahlqvist JR, Widholm P, Leinhard OD, Vissing J. MRI in neuromuscular diseases: an emerging diagnostic tool and biomarker for prognosis and efficacy. *Ann Neurol.* (2020) 88:669–81. doi: 10.1002/ana.25804
29. Díaz-Manera J, Walter G, Straub V. Skeletal muscle magnetic resonance imaging in Pompe disease. *Muscle Nerve*. (2020). doi: 10.1002/mus.27099. [Epub ahead of print].
30. Figueroa-Bonaparte S, Llauger J, Segovia S, Belmonte I, Pedrosa I, Montiel E, et al. Quantitative muscle MRI to follow up late onset Pompe patients: a prospective study. *Sci Rep.* (2018) 8:10898. doi: 10.1038/s41598-018-29170-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 16 March 2021
doi: 10.3389/fneur.2021.634766

Magnetization Transfer Ratio in Lower Limbs of Late Onset Pompe Patients Correlates With Intramuscular Fat Fraction and Muscle Function Tests

Claudia Nuñez-Peralta[1,2], Paula Montesinos[3], Alicia Alonso-Jiménez[4], Jorge Alonso-Pérez[5], David Reyes-Leiva[5], Javier Sánchez-González[3], Jaume Llauger-Roselló[1], Sonia Segovia[5,6], Izaskun Belmonte[7], Irene Pedrosa[7], Antonio Martínez-Noguera[1], Briano Matellini-Mosca[1], Glenn Walter[8] and Jordi Díaz-Manera[5,6,9]**

[1] Radiology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [2] Departament de Medicina, Universidad Autónoma de Barcelona, Barcelona, Spain, [3] Philips Healthcare Iberia, Madrid, Spain, [4] Neuromuscular Reference Center, Neurology Department, University Hospital of Antwerp, Edegem, Belgium, [5] Neuromuscular Disorders Unit, Neurology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [6] Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Madrid, Spain, [7] Rehabilitation Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [8] Department of Physiology and Functional Genomics, University of Florida, Gainesville, FL, United States, [9] John Walton Muscular Dystrophy Research Center, Newcastle University, Newcastle upon Tyne, United Kingdom

Objectives: Magnetization transfer (MT) imaging exploits the interaction between bulk water protons and protons contained in macromolecules to induce signal changes through a special radiofrequency pulse. MT detects muscle damage in patients with neuromuscular conditions, such as limb-girdle muscular dystrophies or Charcot-Marie-Tooth disease, which are characterized by progressive fiber loss and replacement by fatty tissue. In Pompe disease, in which there is, in addition, an accumulation of glycogen inside the muscle fibers, MT has not been tested yet. Our aim is to estimate MT ratio (MTR) in the skeletal muscle of these patients and correlate it with intramuscular fat fraction (FF) and results of muscle function tests.

Methods: We obtained two-point axial Dixon and Dixon-MT sequences of the right thigh on a 1.5 Teslas MRI scanner in 60 individuals, including 29 late onset Pompe disease patients, 2 patients with McArdle disease, and 29 age and sex matched healthy controls. FF and MTR were estimated. Muscle function using several muscle function tests, including quantification of muscle strength, timed test quality of life scales, conventional spirometry obtaining forced vital capacity while sitting and in the supine position, were assessed in all patients.

Results: MTR was significantly lower in Pompe patients compared with controls (45.5 ± 8.5 vs. 51.7 ± 2.3, Student T-test, $p < 0.05$). There was a negative correlation between the MTR and FF muscles studied (correlation coefficient: −0.65, Spearman test: $p < 0.05$). MTR correlated with most of the muscle function test results. We analyzed if there was any difference in MTR values between Pompe patients and healthy controls in those muscles that did not have an increase in fat, a measure that could be related to the

OPEN ACCESS

Edited by:
Edoardo Malfatti,
INSERM U1179 Handicap Neuromusculaire: Physiopathologie, Biothérapie et Pharmacologie Appliquées (END-ICAP), France

Reviewed by:
Marianne De Visser,
University of Amsterdam, Netherlands
Carmelo Rodolico,
University of Messina, Italy

***Correspondence:**
Claudia Nuñez-Peralta
cnunez@santpau.cat
Jordi Díaz-Manera
jordi.diaz-manera@newcastle.ac.uk

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: 28 November 2020
Accepted: 01 February 2021
Published: 16 March 2021

Citation:
Nuñez-Peralta C, Montesinos P, Alonso-Jiménez A, Alonso-Pérez J, Reyes-Leiva D, Sánchez-González J, Llauger-Roselló J, Segovia S, Belmonte I, Pedrosa I, Martínez-Noguera A, Matellini-Mosca B, Walter G and Díaz-Manera J (2021) Magnetization Transfer Ratio in Lower Limbs of Late Onset Pompe Patients Correlates With Intramuscular Fat Fraction and Muscle Function Tests. Front. Neurol. 12:634766. doi: 10.3389/fneur.2021.634766

presence of glycogen in skeletal muscles, but we did not identify significant differences except in the adductor magnus muscle (48.4 ± 3.6 in Pompe vs. 51 ± 1.3 in healthy controls, Student *T*-test = 0.023).

Conclusions: MTR is a sensitive tool to identify muscle loss in patients with Pompe disease and shows a good correlation with muscle function tests. Therefore, the MT technique can be useful in monitoring muscle degeneration in Pompe disease in clinical trials or natural history studies.

Keywords: late onset Pompe disease, lower limb muscle, magnetic transfer ratio, intramuscular fat fraction, muscle function tests

INTRODUCTION

Magnetization transfer (MT) imaging is a magnetic resonance (MRI) technique that exploits the magnetization exchange between water and tissue protons present in different environments. In normal tissue, hydrogen protons are present in two compartments: the so-called "free pool" with mobile protons in free water and a second compartment called the "bound pool," consisting of protons bound to proteins and other macromolecules, such as glycogen (1). Protons within the free pool are responsible for the conventional MRI signal because they have a long and easily detectable T2 signal; on the other hand, protons bound to macromolecules have short T2 values due to their highly restricted motion and are not detected in routine MRI.

The interaction between these two compartments can be probed by measuring the exchange of energy from the bound to the free pool of protons. This is achieved by applying a special radiofrequency (RF) pulse called MT pulse (a low-power RF saturation pulse) that gives origin to the MT effect (2). The MT effect can be used to create additional contrast (MTC) in different tissues, and it is widely used in MR angiography, enhancing the representation of smaller peripheral branches of the vessels (3), or in multiple sclerosis in which the background suppression improves detection of acute lesions (4). The MT effect can also be quantified obtaining the magnetization transfer ratio (MTR), which provides insight into relaxation and exchange rates of free water and macromolecules (5).

Musculoskeletal tissue displays a pronounced MT effect, and it is already demonstrated to be a sensitive measure of muscle damage in patients with neuromuscular conditions, such as limb-girdle muscular dystrophies (LGMD) or Charcot-Marie-Tooth disease, which are disorders characterized by progressive muscle fiber loss and replacement by fatty tissue (6, 7). Nevertheless, the utility of MT has not been tested in Pompe disease yet, a disease in which there is an accumulation of glycogen within muscle fibers leading to cell death and replacement by fat. Estimation of the intramuscular fat fraction (FF) using the Dixon technique in Pompe patients is a useful biomarker for the follow-up of patients as we demonstrate that there is a progressive accumulation of fat in the skeletal muscles that precedes changes in muscle function tests in patients that have already started enzyme replacement therapy (8). However, presymptomatic patients do not always show an increase in fat accumulation even if they develop muscle weakness, suggesting that the identification of changes related to glycogen accumulation could be useful to monitor progression of the disease, especially in these presymptomatic cases. Our hypothesis is that MT could identify early structural changes related to the accumulation of glycogen before degeneration of the muscle fibers to substitution by fatty tissue takes place in the muscles. To confirm our hypothesis, we studied a cohort of patients with late-onset Pompe disease (LOPD) using the MT technique and compared the results with those obtained in healthy controls. Additionally, we correlated MTR values with intramuscular FF and muscle function tests.

METHODS

Study Design and Participants

This is a transversal cross-sectional study involving 60 individuals who were studied at the Hospital de la Santa Creu i Sant Pau (HSCSP) in Barcelona. The study is registered in *ClinicalTrials.gov* with the identifier NCT01914536. The HSCSP ethics committee approved the study, and all participants signed an informed consent form. All study procedures were performed in accordance with Spanish regulations.

The inclusion criteria for the study were (1) genetically confirmed diagnosis of Pompe disease or McArdle disease; (2) no contraindications to MRI; and (3) willingness to complete all muscle function tests, respiratory assessment, and patient-reported outcomes measures. As control, we included a group of healthy volunteers age- and sex-matched with the Pompe patients. We ruled out other neuromuscular conditions in all participants in the study based on clinical data and results of complementary tests, such as blood analysis, spinal MRI, or EMG when needed.

All patients were studied by three physiotherapists with long experience in neuromuscular disorders at HSCSP. The physiotherapists evaluated muscle function using the following tests: the 6-min walking test, time to walk 10 m, timed up-and-go test, time to climb up and down four steps, and the Motor Function Measure 20-item scale (MFM-20). All timed tests were performed asking the patient to not use aids for walking. Muscle strength was studied using both the Medical Research Council (MRC) scale and hand-held myometry. Daily life activities were studied using the activity limitations scale

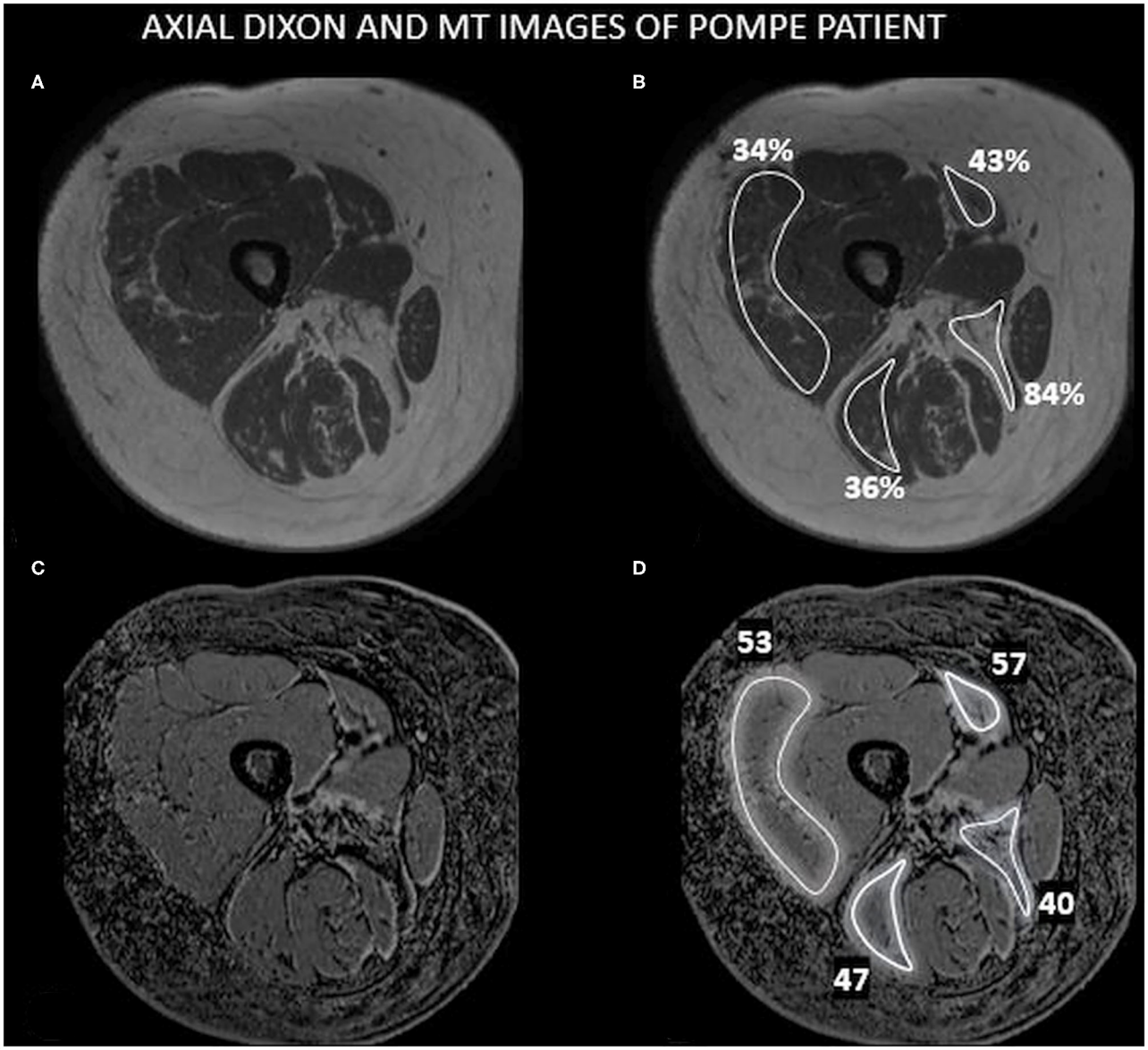


FIGURE 1 | An example of Dixon and MT images. This figure shows an axial image of the right medial thigh of a Pompe patient including the muscles studied in this paper with the ROIs drawn and the values obtained for FF and MTR. **(A)** 3-D Dixon imaging of the thigh. **(B)** ROIs drawn in **(A)** showing the results of FF observed. **(C)** Dixon-MTC imaging of the thigh. **(D)** ROIs drawn in **(C)** showing the results of MTR.

for patients with upper and/or lower limb impairments, and quality of life was analyzed using both the Individualized Neuromuscular Quality of Life Questionnaire and the Short Form 36 questionnaire. We obtained forced vital capacity (FVC), both seated and lying down, using the Carefusion Microlab ML 3500 MK8 spirometer (Carefusion, Yorba Linda, CA, USA). These last two tests were added due their being commonly used in Pompe disease patients to measure their clinical status (9). Finally, Creatine kinase levels in serum were quantified in the HSCSP biochemistry laboratory following standard protocols.

MRI

All patients were examined in a 1.5 MR system (1.5 Achieva dStream; Philips, Best, NL) at HSCSP in the supine position with stretched legs using a 32-channel body coil. Axial 3-D Dixon FFE on the middle third of both thighs was performed with the following parameters: TR/TE1/TE2 = 5.78/1.8/4 ms, flip angle = 15, FOV = 520 × 340 × 300, acquired voxel size = 1 × 1 × 3 mm^3; acquisition time was 2 min 33 s. Also, another 3-D Dixon FFE sequence was acquired with and without an off-resonance magnetization transfer contrast prepulse active, TR/TE1/TE2 = 32/1.8/4 ms, flip angle = 15, FOV = 400 × 200 × 114 mm,

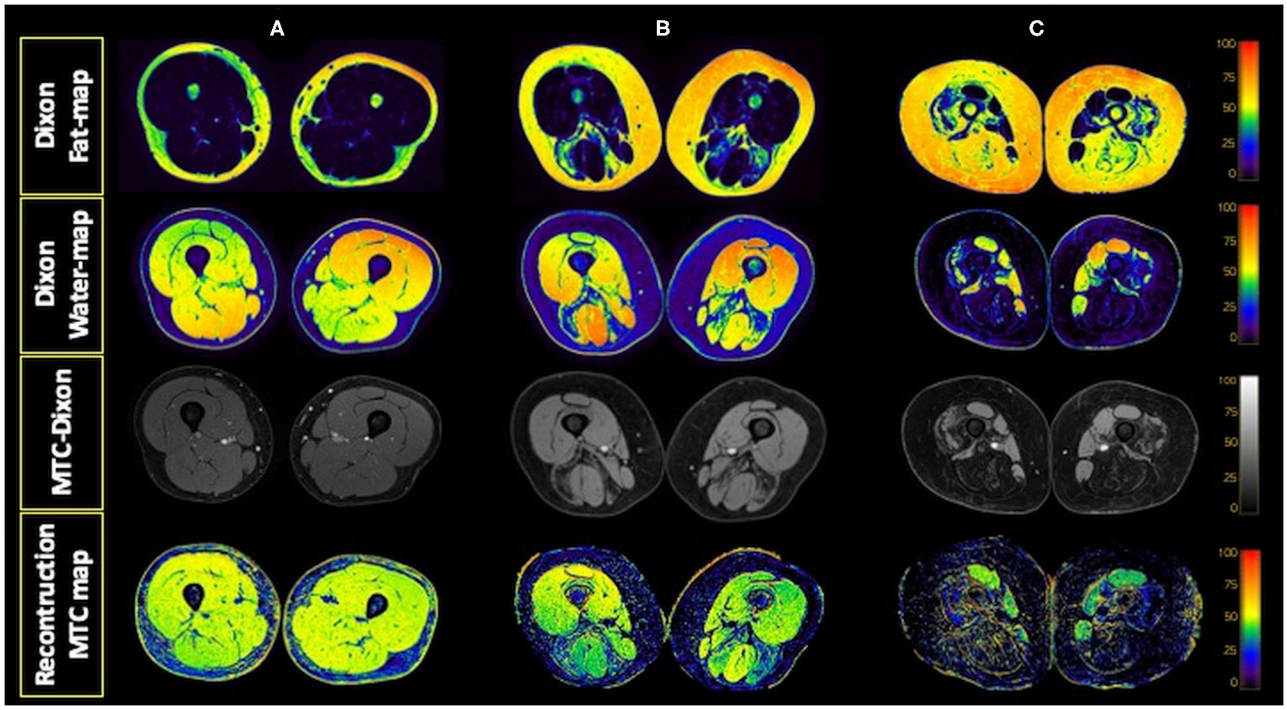


FIGURE 2 | Dixon water and fat maps and MTC-Dixon images obtained. Examples of fat and water maps generated from the data obtained with the Dixon sequence and of the images obtained with the MTC-Dixon sequence and the map generated. Axial sections of the thighs of a control **(A)** and two patients with Pompe disease, one with moderate involvement **(B)** and the other with severe muscle involvement **(C)**.

TABLE 1 | Mean demographic and clinical data of the patients included in the study.

	Pompe	McArdle	Controls
Individuals (*n*)	29	2	29
Age at MRI	41.2 ± 21.3 y.o.	44 and 38 y.o	45.2 ± 22.7 y.o.
Use of ERT (*n*)	21		
Aids for walking (*n*)	12	–	–
Need of ventilation (*n*)	11	–	–

MRI, magnetic resonance imaging; ERT, enzymatic replacement therapy; N, number of patients.

acquired voxel size = 1 × 1 × 3 mm^3; total acquisition time for images with and without MTC module active was 7 min 58 s.

Analysis

We analyzed both sequences in the middle right thigh of all patients. Regions of interest (ROIs) were manually drawn by three investigators on one slice of the following muscles: vastus lateralis (VL), sartorius (Sar), the long head of the biceps femoris (BLH), and adductor magnus (AM), keeping a reasonable distance from the fascia and subcutaneous fat tissue. FF was obtained using a Philips Research Image Development Environment (PRIDE) tool developed for this purpose. MTR was calculated using the following formula in which Mo and MSat refer to the images without and with the saturation prepulse. MTR values were expressed in percentage. **Figure 1** shows an example of MT imaging, and the ROIs drawn and **Figure 2** shows an example of the images obtained in this study.

$$MTR = \left(\frac{Mo - M\,Sat}{Mo} \right) \quad (1)$$

Statistics

We confirmed that the variables to be analyzed were normally distributed using a Kolmogorov–Smirnov test, and therefore, parametric statistical studies were used. We used the Student *T*-test to analyze if the differences observed in FF and MTR between Pompe and controls were statistically significant. We used the Pearson test to study if correlations between variables were significant. We consider that a correlation was good if the correlation coefficient was higher than 0.6. The significance level for all statistical studies was set at $p < 0.05$, and *post-hoc* Bonferroni corrections were used when needed. Statistical studies were performed with SPSS for Mac computers (Version 21, SPSS Inc., Chicago, IL).

RESULTS

Patients Included

We included in the study 29 LOPD patients of which 23 patients were symptomatic and treated with enzymatic replacement therapy (ERT) and 6 patients were presymptomatic and not

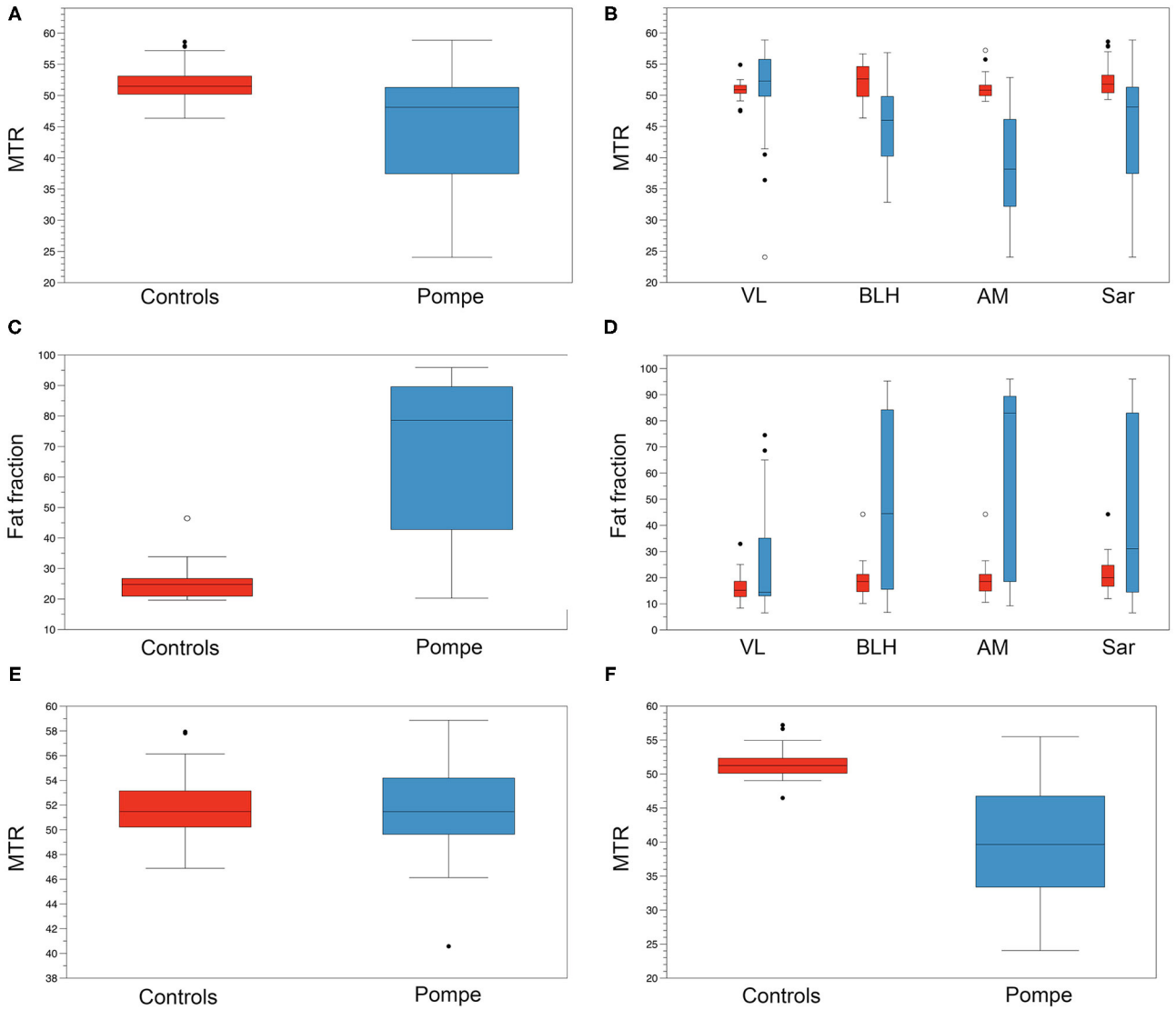


FIGURE 3 | MTR and FF in controls and Pompe patients. **(A)** Mean thigh MTR value in controls (red) and Pompe patients (blue). **(B)** MTR value of VL, BLH, AM, and Sar in controls (red) and Pompe patients (blue). **(C)** Mean thigh FF value in controls (red) and Pompe patients (blue). **(D)** FF of VL, BLH, AM, and Sar in controls (red) and Pompe patients (blue). **(E)** Mean thigh MTR value in individuals with mean thigh FF lower than 20% in controls (red) and Pompe patients (blue). **(F)** Mean thigh MTR value in individuals with mean thigh FF higher than 20% in controls (red) and Pompe patients (blue). Mean value and standard deviation are shown.

treated with ERT. Clinical features of LOPD patients included in this study have been reported before (10). We also included in the study 2 patients with McArdle disease without muscle weakness at clinical examination and 29 healthy age- and sex-matched controls. Main demographic and clinical data of the individuals included in the study are summarized in **Table 1**.

MRI Results

We found lower MTR average value in LOPD patients' muscles (45.5 ± 8.5) compared with controls (51.7 ± 2.3) (Student *T*-test, $p < 0.001$; **Figure 3A**). When analyzed separately, we identified significant lower values in AM and BLH but not in VL nor Sar in LOPD patients compared with controls (**Figure 3B**). MTR average value was lower in symptomatic compared with presymptomatic (41.9 ± 9.8 vs. 51 ± 4, respectively, Student *T*-test $p < 0.001$). There were no differences in MTR between controls and presymptomatic patients.

We found higher mean thigh FF in LOPD patients' muscles (44.9 ± 32.8%) compared with controls (17.86 ± 6.8%) (Student *T*-test, p <0.001; **Figure 3C**). When analyzed separately, we identified higher values in AM, BLH, and VL but not in Sar in LOPD patients compared with controls (**Figure 3D**). Mean thigh FF was significantly higher in symptomatic compared with presymptomatic LOPD patients (32.6 ± 7.7 vs. 21.24 ± 7.3, respectively, Student *T*-test $p < 0.001$). There were no differences in mean thigh FF between controls and presymptomatic patients.

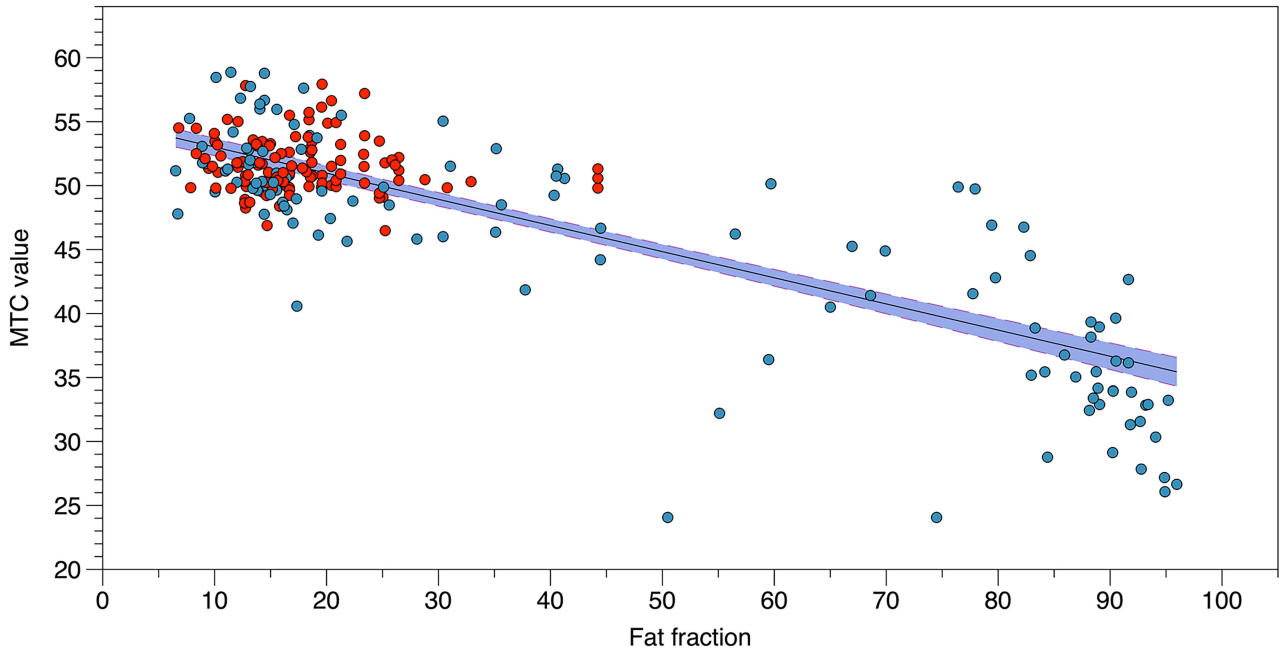


FIGURE 4 | Correlation between MTR value and FF. The figure shows the correlation between MTR value and FF in each muscle analyzed from controls (red) and patients with Pompe disease (blue).

We did not identify significant differences in MTR in those muscles with FF lower than 20% between LOPD and healthy controls when the four individual muscles were analyzed together (**Figures 3E,F**). However, when analyzed separately, we observed significant differences in MTR in the AM (48.4 ± 3.6 in Pompe vs. 51 ± 1.3 in controls, Student *T*-test = 0.023). We observed a negative correlation between MTR value and FF in muscles studied (correlation coefficient: −0.65, Spearman test $p < 0.0001$; **Figure 4**).

Correlation With Results of the Muscle Function Tests

We observed a significant correlation between mean thigh MTR value and the results of muscle function tests, including timed tests, such as the 6MWT or the time to walk 10 m, and measures of muscle strength, such as the composite MT value of the lower limbs. Moreover, we identified a good correlation between individual muscle MTR values and specific muscle assessments as is shown in **Table 2**.

TABLE 2 | Correlation between MTC values and results of the muscle function tests.

	Mean MTC value	
	Correlation coefficient	**Significance**
Time to walk 10 m	−0.72	0.0001
6 MWT	0.813	0.0001
Time up go	−0.78	0.0001
Time climb up 4 Steps	−0.79	0.0001
Time go down 4 Steps	−0.75	0.0001
MRC lower limbs	0.71	0.0001
Activlim	0.59	0.001
CK levels	0.12	0.71
MTC adductor magnus		
Hip adduction	0.63	0.0001
MTC biceps long head		
Knee flexion	0.54	0.003
MTC vastus lateralis		
Knee extension	0.38	0.52

6MWT, 6 min walking Test. Pearson test was performed, and p-values are shown.

DISCUSSION

We observed that MT is an indirect measurement of muscle loss in LOPD patients, and it correlates with results of different muscle function tests commonly used in clinical trials and natural history studies. We also identified a good correlation between muscle FF and muscle MTR value, suggesting that MTR values decrease in relation to progressive loss of muscle fibers that are replaced by fatty tissue.

Quantitative muscle MRI is progressively being implemented in the follow-up of patients with neuromuscular diseases, including Pompe disease, in clinical trials and natural history studies (11–13). Dixon imaging allows calculating the FF, which is the amount of skeletal muscle replaced by fatty tissue. FF is shown to correlate with the results of muscle function tests in different diseases, including but not limited to Pompe disease.

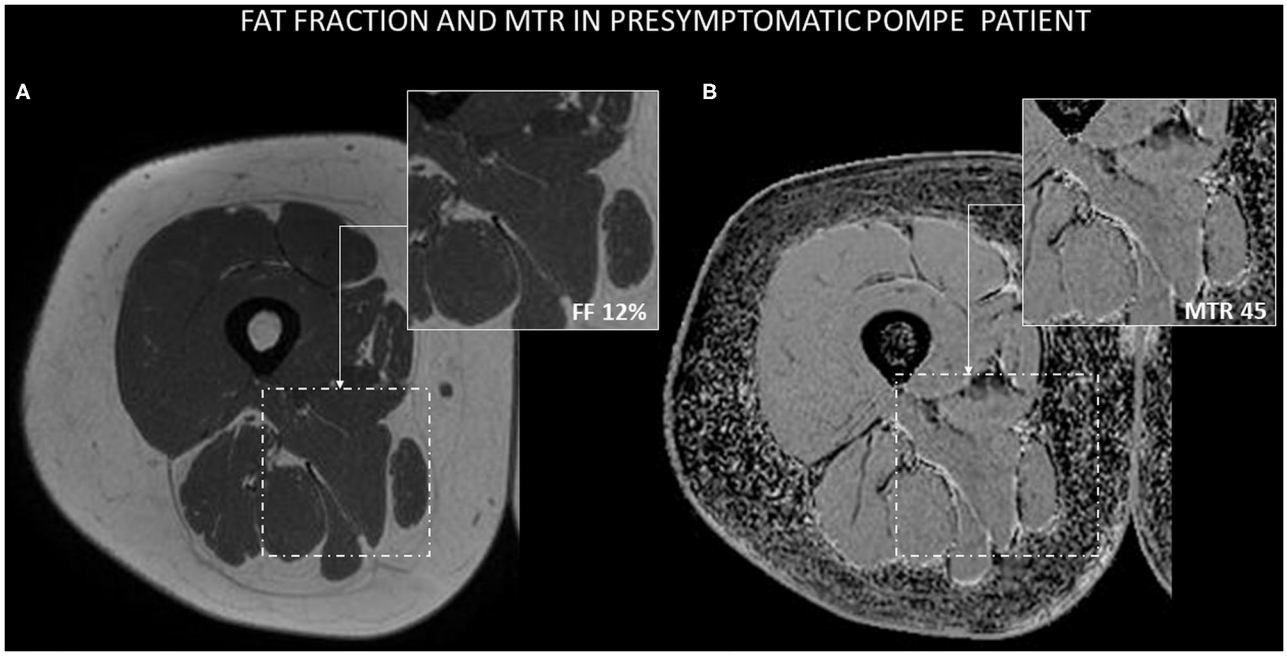


FIGURE 5 | Example of Dixon and MTR images in a presymptomatic Pompe patient. **(A)** AM muscle is not macroscopically infiltrated by fat, and the estimation of FF using the Dixon technique is 12%. **(B)** MTR value in AM is 45.

Moreover, Dixon has proved to be useful to follow up the progression of fat replacement in LOPD patients in longitudinal studies. Previous studies have identified risk factors associated with FF, such as early age of onset of symptoms and disease duration (8).

Based on these results, Dixon sequences are being implemented in the new clinical trials that are being designed in this disease. MT imaging is based on the magnetization exchange between protons present in the free-water compartment and protons bound to macromolecules. The more hydrated the macromolecules are, the more magnetization transfer is obtained. MT from fat is low as it is a hydrophobic tissue. Hydrophobic properties of lipids cause them to experience limited MT; for that reason, MT is generally low in tissues with high fat content (14–16). However, it is known that fat generates chemical shift artifacts that could compromise the characterization and analysis of tissue. The use of traditional fat-suppression techniques in MT studies can affect the MTR measurements, but the use of water images obtained from Dixon sequences avoids possible confounding effects derived from the fat signal for MTR calculation (17, 18). For these reasons, in this study, we decided to use a two-point Dixon sequence with an off-resonance prepulse. However, the signal obtained in muscles replaced by fat could also be influenced by other concomitant components, such as fibrotic tissue or inflammation. Therefore, MTR is not directly measuring only fat replacement, but in our opinion, is an indirect measurement of loss of muscle fibers that reduce the values obtained. In this sense, our results are similar to the ones reported in other neuromuscular diseases, such as LGMD, inclusion body myositis, Charcot-Marie-Tooth disease, or spinal muscle atrophy (6, 7). We observed lower MTR values in symptomatic patients compared with non-symptomatic patients and controls. To fully understand the biophysical origins of MT in Pompe muscle, a more detailed quantitative MT study is warranted that takes into account pool sizes, T1, lipids, and multiple transfer magnetization transfer mechanisms (5, 19).

Our initial hypothesis was that MT could be useful as an indirect measurement of glycogen. Glycogen is a highly hydrated molecule, and therefore, it was tempting to hypothesize that, in patients with Pompe disease who accumulate glycogen in their muscles, MTR could increase because of the accumulation of water bound to the glycogen. In this study, we observed that fat replacement is a confounding factor because it reduces the MTR signal considerably and makes difficult the identification of the effect of other tissue components on MTR. However, MT values in muscles with low levels of fat replacement (<20% FF) are not statistically different between healthy controls and LOPD patients except for the AM. Additionally, we have not seen differences between presymptomatic non-treated and symptomatic treated patients in the MTR values in muscles with FF lower than 20% (**Figure 5**). There are two potential explanations for our findings. On one hand, it is possible that the amount of glycogen accumulated in these muscles was too low to induce any change in MTR. It is well-known that LOPD patients accumulate less glycogen than infantile onset patients (IOPD). Muscle biopsies of LODP patients can show mild accumulation of glycogen or even be normal, and muscle biopsies in IOPD patients are characterized by a massive accumulation of glycogen. On the other hand and based on the physiopathology of the disease, it is probable that

the accumulation of glycogen was more prominent in those muscles undergoing fatty replacement, but if that was the case, MTR could be influenced by the loss of muscle fibers and expansion of fat, producing, as a result, a decrease in MTR values (20–22). We have also tested MT in 2 patients with glucogenosis type V, usually known as McArdle disease, who had mean thigh FF lower than 10%, and we have not seen any change in MTR values compared with controls. In our opinion, to completely rule out MTR as an indirect measure of glycogen in patients, we should explore it in pretreated IOPD patients because they accumulate larger amounts of glycogen in their muscles.

Interestingly, we have seen a statistically significant decrease in MTR values in the AM muscle in presymptomatic patients without an increase in FF. We previously identified that this muscle is one of the muscles earlier replaced by fat in the progression of the disease (8). Two of these patients developed mild hip adduction weakness soon after being scanned in this study. In our opinion, this result probably reflects the loss of muscle fibers that could be the cause of the weakness rather than accumulation of glycogen as we would expect an increase in the MTR values. It is possible that MT has high sensitivity to detect mild muscle fiber loss even when these fibers have not been replaced by fat yet, and Dixon studies do not detect an increase in fatty tissue. Accordingly, McDaniel et al. reported a decrease in MTR value in muscles of patients with muscular dystrophies that were not replaced by fat, which could be related, in our opinion, to the existence of early muscle damage not yet leading to fat replacement.

The main limitation of our study is that we have not used a gold standard test to study the amount of glycogen present in the muscles of the patients, such as a muscle biopsy or specific imaging sequences able to identify glycogen, such as 13C and 1H spectroscopy (23), chemical exchange saturation transfer imaging of glycogen (GlycoCEST) (24–26), and the recently described nuclear Overhauser enhancement of glycogen (Glyco-nOe) (27, 28). All these sequences are able to identify glycogen in muscles and/or liver but are not available in conventional hospitals not focused on research. Some of the advantages of the MT technique are that it is available in many hospitals, does not require a specific coil, and the acquisition is relatively fast and easy to analyze.

We have seen that MT values correlate with FF and, therefore, could be used to monitor disease progression over time although this should be explored in future longitudinal studies. The only longitudinal study published so far shows that MT measurements decreased in Charcot-Marie-Tooth and IBM patients after 1 year of follow-up in relation to the accumulation of fatty tissue identified using Dixon (29).

CONCLUSION

The MTR estimation constitutes a sensitive tool for the identification of cellular damage in patients with Pompe disease. It is correlated with muscle loss and muscle function tests. Therefore, MT should be further explored as a tool for monitoring muscle degeneration in Pompe disease in clinical trials or natural history studies.

DATA AVAILABILITY STATEMENT

The original contributions generated for the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Hospital de la Santa Creu i Sant Pau, Barcelona. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

CN-P and PM: concept and design of the study, acquisition and analysis of data, and drafting the paper. AA-J and JA-P: acquisition and analysis of data, and drafting the paper. DR-L, JL-R, IB, IP, and BM-M: concept and design of the study, acquisition of data, and drafting the paper. JS-G, AM-N and GW: concept and design of the study, and drafting the paper. SS: concept and design of the study, coordination of the study, and drafting the paper. JD-M: concept and design of the study, obtaining funding for the study, acquisition and analysis of data, and drafting the paper. All authors contributed to the article and approved the submitted version.

FUNDING

This study was sponsored by grants from Sanofi Genzyme and from the Spanish Ministry of Health, Fondos FEDER-ISCIII PI18/01525 to JD-M. The company has neither reviewed the data nor suggested changes in any of the conclusions of the paper.

ACKNOWLEDGMENTS

We would like to thank all members of the Spanish Association of Patients with Glycogenosis (www.glucogenosis.org) and the Philips Healthcare engineer team for their collaboration. We also thank all patients and relatives for their collaboration in this study.

REFERENCES

1. Sinclair CDJ, Samson RS, Thomas DL, Weiskopf N, Lutti A, Thornton JS, et al. Quantitative magnetization transfer in *in vivo* healthy human skeletal muscle at 3 T. *Magn Reson Med.* (2010) 64:1739–48. doi: 10.1002/mrm.22562
2. Henkelman RM, Stanisz GJ, Graham SJ. Magnetization transfer in MRI: a review. *NMR Biomed.* (2001) 14:57–64. doi: 10.1002/nbm.683
3. Catalano C, Pavone P, Laghi A, Faroni J, Clementi M, Di Girolamo M, et al. Magnetic resonance angiography of the intracranial circle using magnetization transfer contrast (MTC). *Radiol Med.* (1995) 89:245.

4. Horsfield MA. Magnetization transfer imaging in multiple sclerosis. *J Neuroimaging.* (2005) 15(Suppl. 4):58–67. doi: 10.1177/1051228405282242
5. Li K, Dortch RD, Kroop SF, Huston JW, Gochberg DF, Park JH, et al. A rapid approach for quantitative magnetization transfer imaging in thigh muscles using the pulsed saturation method. *Magn Reson Imaging.* (2015). 33:709–17. doi: 10.1016/j.mri.2015.03.003
6. Kollmer J, Kessler T, Sam G, Hayes JM, Lentz SI, Heiland S, et al. Magnetization transfer ratio: a quantitative imaging biomarker for 5q spinal muscular atrophy. *Eur J Neurol.* (2020) 28:331–40. doi: 10.1111/ene.14528
7. McDaniel JD, Ulmer JL, Prost RW, Franczak MB, Jaradeh S, Hamilton CA, et al. Magnetization transfer imaging of skeletal muscle in autosomal recessive limb girdle muscular dystrophy. *J Comput Assist Tomogr.* (1999) 23:609–14. doi: 10.1097/00004728-199907000-00023
8. Nuñez-Peralta C, Alonso-Pérez J, Llauger J, Segovia S, Montesinos P, Belmonte I, et al. Follow-up of late-onset Pompe disease patients with muscle magnetic resonance imaging reveals increase in fat replacement in skeletal muscles. *J Cachexia Sarcopenia Muscle.* (2020) 11:1032–46. doi: 10.1002/jcsm.12555
9. Berger KI, Kanters S, Jansen JP, Stewart A, Sparks S, Haack KA, et al. Forced vital capacity and cross-domain late-onset Pompe disease outcomes: an individual patient-level data meta-analysis. *J Neurol.* (2019) 266:2312–21. doi: 10.1007/s00415-019-09401-1
10. Figueroa-Bonaparte S, Llauger J, Segovia S, Belmonte I, Pedrosa I, Montiel E, et al. Quantitative muscle MRI to follow up late onset Pompe patients: a prospective study. *Sci Rep.* (2018) 8:10898. doi: 10.1038/s41598-018-29170-7
11. Paoletti M, Pichiecchio A, Piccinelli SC, Tasca G, Berardinelli AL, Padovani A, et al. Advances in quantitative imaging of genetic and acquired myopathies: clinical applications and perspectives. *Front Neurol.* (2019) 10:78. doi: 10.3389/fneur.2019.00078
12. Díaz-Manera J, Walter G, Straub V. Skeletal muscle magnetic resonance imaging in Pompe disease. *Muscle Nerve.* (2020). doi: 10.1002/mus.27099. [Epub ahead of print].
13. Nuñez-Peralta C, Alonso-Pérez J, Díaz-Manera J. The increasing role of muscle MRI to monitor changes over time in untreated and treated muscle diseases. *Curr Opin Neurol.* (2020) 33:611–20. doi: 10.1097/WCO.0000000000000851
14. Kim H, Booth CJ, Pinus AB, Chen P, Lee A, Qiu M, et al. Induced hepatic fibrosis in rats: hepatic steatosis, purpose: methods: results: conclusion: distribution. *Radiology.* (2008). 247:696–705. doi: 10.1148/radiol.2473070605
15. Shah B, Anderson SW, Scalera J, Jara H, Soto JA. Quantitative MR imaging: physical principles and sequence design in abdominal imaging. *Radiographics.* (2011) 31:867–80. doi: 10.1148/rg.313105155
16. Yarnykh VL, Tartaglione EV, Ioannou GN. Fast macromolecular proton fraction mapping of the human liver *in vivo* for quantitative assessment of hepatic fibrosis. *NMR Biomed.* (2015) 28:1716–25. doi: 10.1002/nbm.3437
17. Smith AK, Dortch RD, Dethrage LM, Lyttle BD, Kang H, Welch EB, et al. Incorporating dixon multi-echo fat water separation for novel quantitative magnetization transfer of the human optic nerve *in vivo*. *Magn Reson Med.* (2017) 77:707–16. doi: 10.1002/mrm.26164
18. Ercan E, Varma G, Dimitrov IE, Xi Y, Pinho MC, Yu FF, et al. Combining inhomogeneous magnetization transfer and multipoint Dixon acquisition: potential utility and evaluation. *Magn Reson Med.* (2020) 85:2136–44. doi: 10.1002/mrm.28571
19. Zhou Y, van Zijl PCM, Xu J, Yadav NN. Mechanism and quantitative assessment of saturation transfer for water-based detection of the aliphatic protons in carbohydrate polymers. *Magn Reson Med.* (2020) 85:1643–54. doi: 10.1002/mrm.28503
20. Sinclair CDJ, Morrow JM, Miranda MA, Davagnanam I, Cowley PC, Mehta H, et al. Skeletal muscle MRI magnetisation transfer ratio reflects clinical severity in peripheral neuropathies. *J Neurol Neurosurg Psychiatry.* (2012) 83:29–32. doi: 10.1136/jnnp.2011.246116
21. Brown JWL, Chowdhury A, Kanber B, Prados Carrasco F, Eshaghi A, Sudre CH, et al. Magnetisation transfer ratio abnormalities in primary and secondary progressive multiple sclerosis. *Mult Scler J.* (2019)26:679–87. doi: 10.1177/1352458519841810
22. Schmierer K, Scaravilli F, Altmann DR, Barker GJ, Miller DH. Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. *Ann Neurol.* (2004) 56:407–15. doi: 10.1002/ana.20202
23. Heinicke K, Dimitrov IE, Romain N, Cheshkov S, Ren J, Malloy CR, et al. Reproducibility and absolute quantification of muscle glycogen in patients with glycogen storage disease by C NMR spectroscopy at 7 Tesla. *PLoS ONE.* (2014). 9:e108706. doi: 10.1371/journal.pone.0108706
24. Zijl PCM Van, Jones CK, Ren J, Malloy CR, Sherry AD. MRI detection of glycogen *in vivo* by using chemical exchange saturation transfer imaging (glycoCEST). *Proc Natl Acad Sci USA.* (2007). 104:4359–64. doi: 10.1073/pnas.0700281104
25. Stender S, Zaha VG, Malloy CR, Sudderth J, DeBerardinis RJ, Park JM. Assessment of rapid hepatic glycogen synthesis in humans using dynamic 13C magnetic resonance spectroscopy. *Hepatol Commun.* (2020) 4:425–33. doi: 10.1002/hep4.1458
26. Miller CO, Cao J, Chekmenev EY, Damon BM, Cherrington AD, Gore JC. Noninvasive measurements of glycogen in perfused mouse livers using chemical exchange saturation transfer NMR and comparison to (13)C NMR spectroscopy. *Anal Chem.* (2015) 87:5824–30. doi: 10.1021/acs.analchem.5b01296
27. Zhou Y, van Zijl PCM, Xu X, Xu J, Li Y, Chen L, et al. Magnetic resonance imaging of glycogen using its magnetic coupling with water. *Proc Natl Acad Sci USA.* (2020) 117:3144–9. doi: 10.1073/pnas.1909921117
28. Vinogradov E. MRI monitoring of energy storage *in vivo* using magnetization pathways. *Proc Natl Acad Sci USA.* (2020) 117:5092–94. doi: 10.1073/pnas.2001048117
29. Morrow JM, Sinclair CD, Fischmann A, Machado PM, Reilly MM, Yousry TA, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. *Lancet Neurol.* (2016) 15:65–77. doi: 10.1016/S1474-4422(15)00242-2

Conflict of Interest: PM was employed by company Philips Healthcare Iberia.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

CORRECTION
published: 13 July 2021
doi: 10.3389/fneur.2021.727020

Corrigendum: Magnetization Transfer Ratio in Lower Limbs of Late Onset Pompe Patients Correlates With Intramuscular Fat Fraction and Muscle Function Tests

Claudia Nuñez-Peralta[1,2], Paula Montesinos[3], Alicia Alonso-Jiménez[4], Jorge Alonso-Pérez[5], David Reyes-Leiva[5], Javier Sánchez-González[3], Jaume Llauger-Roselló[1], Sonia Segovia[5,6], Izaskun Belmonte[7], Irene Pedrosa[7], Antonio Martínez-Noguera[1], Briano Matellini-Mosca[1], Glenn Walter[8] and Jordi Díaz-Manera[5,6,9]**

[1] Radiology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [2] Departament de Medicina, Universidad Autónoma de Barcelona, Barcelona, Spain, [3] Philips Healthcare Iberia, Madrid, Spain, [4] Neuromuscular Reference Center, Neurology Department, University Hospital of Antwerp, Edegem, Belgium, [5] Neuromuscular Disorders Unit, Neurology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [6] Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Madrid, Spain, [7] Rehabilitation Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain, [8] Department of Physiology and Functional Genomics, University of Florida, Gainesville, FL, United States, [9] John Walton Muscular Dystrophy Research Center, Newcastle University, Newcastle upon Tyne, United Kingdom

Keywords: late onset Pompe disease, lower limb muscle, magnetic transfer ratio, intramuscular fat fraction, muscle function tests

OPEN ACCESS

Approved by:
Frontiers Editorial Office, Frontiers Media SA, Switzerland

****Correspondence:***
Claudia Nuñez-Peralta
cnunez@santpau.cat
Jordi Díaz-Manera
jordi.diaz-manera@newcastle.ac.uk

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: *17 June 2021*
Accepted: *18 June 2021*
Published: *13 July 2021*

Citation:
Nuñez-Peralta C, Montesinos P, Alonso-Jiménez A, Alonso-Pérez J, Reyes-Leiva D, Sánchez-González J, Llauger-Roselló J, Segovia S, Belmonte I, Pedrosa I, Martínez-Noguera A, Matellini-Mosca B, Walter G and Díaz-Manera J (2021) Corrigendum: Magnetization Transfer Ratio in Lower Limbs of Late Onset Pompe Patients Correlates With Intramuscular Fat Fraction and Muscle Function Tests. Front. Neurol. 12:727020. doi: 10.3389/fneur.2021.727020

A Corrigendum on

Magnetization Transfer Ratio in Lower Limbs of Late Onset Pompe Patients Correlates With Intramuscular Fat Fraction and Muscle Function Tests

by Nuñez-Peralta C, Montesinos P, Alonso-Jiménez A, Alonso-Pérez J, Reyes-Leiva D, Sánchez-González J, et al. (2021). Front. Neurol. 12:634766. doi: 10.3389/fneur.2021.634766

In the published article, there was an error regarding the affiliation(s) for Claudia Nuñez-Peralta. As well as having affiliation(s) ** 1 **, they should also have **2Departament de Medicina, Universidad Autónoma de Barcelona, Barcelona, Spain**.

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

frontiers in Neurology

ORIGINAL RESEARCH
published: 22 March 2021
doi: 10.3389/fneur.2021.633808

Combination of Quantitative MRI Fat Fraction and Texture Analysis to Evaluate Spastic Muscles of Children With Cerebral Palsy

Tugba Akinci D'Antonoli[1,2], Francesco Santini[3,4], Xeni Deligianni[3,4], Meritxell Garcia Alzamora[2,5], Erich Rutz[6,7], Oliver Bieri[1,3], Reinald Brunner[8,9] and Claudia Weidensteiner[3,4]*

[1] Department of Pediatric Radiology, University Children's Hospital Basel, Basel, Switzerland, [2] Department of Radiology, University Hospital of Basel, Basel, Switzerland, [3] Division of Radiological Physics, Department of Radiology, University Hospital of Basel, Basel, Switzerland, [4] Department of Biomedical Engineering, University of Basel, Basel, Switzerland, [5] Division of Diagnostic and Interventional Neuroradiology, University Hospital of Basel, Basel, Switzerland, [6] Pediatric Orthopedic Department, Murdoch Children's Research Institute, The Royal Children's Hospital, MCRI the University of Melbourne, Melbourne, VIC, Australia, [7] Faculty of Medicine, The University of Basel, Basel, Switzerland, [8] University Children's Hospital Basel, Basel, Switzerland, [9] Department of Orthopedic Surgery, University Children's Hospital Basel, Basel, Switzerland

OPEN ACCESS

Edited by:
Wenbin Guo, Second Xiangya Hospital, Central South University, China

Reviewed by:
Martin Svehlik, Medical University of Graz, Austria
Ronald Peeters, University Hospitals Leuven, Belgium

***Correspondence:**
Tugba Akinci D'Antonoli
tugba.akincidantonoli@usb.ch

Specialty section:
This article was submitted to Applied Neuroimaging, a section of the journal Frontiers in Neurology

Received: *26 November 2020*
Accepted: *01 March 2021*
Published: *22 March 2021*

Citation:
Akinci D'Antonoli T, Santini F, Deligianni X, Garcia Alzamora M, Rutz E, Bieri O, Brunner R and Weidensteiner C (2021) Combination of Quantitative MRI Fat Fraction and Texture Analysis to Evaluate Spastic Muscles of Children With Cerebral Palsy. Front. Neurol. 12:633808. doi: 10.3389/fneur.2021.633808

Background: Cerebral palsy (CP) is the most common cause of physical disability in childhood. Muscle pathologies occur due to spasticity and contractures; therefore, diagnostic imaging to detect pathologies is often required. Imaging has been used to assess torsion or estimate muscle volume, but additional methods for characterizing muscle composition have not thoroughly been investigated. MRI fat fraction (FF) measurement can quantify muscle fat and is often a part of standard imaging in neuromuscular dystrophies. To date, FF has been used to quantify muscle fat and assess function in CP. In this study, we aimed to utilize a radiomics and FF analysis along with the combination of both methods to differentiate affected muscles from healthy ones.

Materials and Methods: A total of 9 patients (age range 8–15 years) with CP and 12 healthy controls (age range 9–16 years) were prospectively enrolled (2018–2020) after ethics committee approval. Multi-echo Dixon acquisition of the calf muscles was used for FF calculation. The images of the second echo (TE = 2.87 ms) were used for feature extraction from the soleus, gastrocnemius medialis, and gastrocnemius lateralis muscles. The least absolute shrinkage and selection operator (LASSO) regression was employed for feature selection. RM, FF model (FFM), and combined model (CM) were built for each calf muscle. The receiver operating characteristic (ROC) curve and their respective area under the curve (AUC) values were used to evaluate model performance.

Results: In total, the affected legs of 9 CP patients and the dominant legs of 12 healthy controls were analyzed. The performance of RM for soleus, gastrocnemius medialis, and gastrocnemius lateralis (AUC 0.92, 0.92, 0.82, respectively) was better than the FFM (AUC 0.88, 0.85, 0.69, respectively). The combination of both models always had a better performance than RM or FFM (AUC 0.95, 0.93, 0.83). FF was higher in the patient group (FF_S 9.1%, FF_{GM} 8.5%, and FF_{GL} 10.2%) than control group (FF_S 3.3%, FF_{GM} 4.1%, FF_{GL} 6.6%).

Conclusion: The combination of MRI quantitative fat fraction analysis and texture analysis of muscles is a promising tool to evaluate muscle pathologies due to CP in a non-invasive manner.

Keywords: cerebral palsy, pediatric imaging, dixon imaging, intramuscular fat, magnetic resonance imaging, radiomics analysis, texture analysis

INTRODUCTION

Cerebral palsy (CP) is the most common cause of physical disability in childhood, caused by brain injury during the antenatal or early postnatal period (1). Although primary damage occurs in the central nervous system, clinical symptoms are mostly associated with the peripheral neuromuscular system, particularly with skeletal muscles (2). The severity of the clinical manifestations depends on the degree of the injury, ranging from mild movement disorder to severe functional limitation (2). Muscle pathologies occur due to spasticity and contractures, and so far, those pathologies are assessed by either clinical scoring systems, e.g., modified Ashworth scale (MAS), or invasive procedures, e.g., biopsies (2). Imaging has been used to assess torsion or estimate muscle volume (3, 4), but additional methods for characterizing muscle composition have not thoroughly been investigated.

Quantitative magnetic resonance imaging (MRI) is a promising non-invasive imaging modality to assess pathologic changes in muscles. Particularly in neuromuscular muscle diseases, quantitative MRI methods have already become standard for disease monitoring (5–8). Among these quantification methods, fat fraction (FF) measurement is commonly employed to determine fatty infiltration in a muscle, providing insights into function and pathophysiology (6, 7, 9, 10). Most of the quantification methods are based on a mean value calculation within a region of interest (ROI); however, mean values cannot entirely capture the heterogeneity or dynamic variations within the ROI and, therefore, will not show a robust correlation with tissue characteristics (11, 12). So far, FF analysis has been rarely employed to evaluate CP patients (13, 14).

Texture analysis (also called radiomics) is an advanced technique that aims to extract quantitative parameters from diagnostic images to discover the relationship between imaging features and the underlying biological information (15). To date, radiomics analysis has been mostly applied in the field of oncology—including but not limited to gene-expression pattern prediction (16), lesion characteristic discrimination (17), and treatment outcome prediction (18). Radiomics analysis of skeletal muscles recently gained more attention with the increased understanding of the relationship between muscle texture changes and disease pathophysiology (19, 20). Up to now, few studies have focused on radiomics analysis of skeletal muscles, and most of them were either animal studies or in healthy populations (21–27). The potential of the texture analysis of pathologic muscles in human subjects has rarely been explored (28, 29). To our knowledge, our study is the first to employ muscle texture analysis along with FF measurement in children with CP.

TABLE 1 | All study participant demographics.

Parameter	Control group ($n = 12$)	Cerebral palsy group ($n = 9$)	Fisher's exact test *P* value
Age	11.1 (9.6–13.7)	11.5 (10.6–12.0)	0.730
Sex			0.061
Female, *n* (%)	6 (50%)	1 (11%)	
Male, *n* (%)	6 (50%)	8 (89%)	
Height	140.0 (134.0–162.5)	146.0 (135.0–151.0)	0.634
Weight	32.35 (28.15–49.25)	32.7 (26.6–44.6)	0.822
BMI	16.4 (15.7–17.3)	16.1 (14.6–18.3)	0.861

Unless otherwise specified, data are medians and interquartile ranges. BMI, body mass index.

In this study, we aimed to employ radiomics and FF analysis along with the combination of both methods to differentiate pathologic muscles from healthy ones in children with CP and healthy controls and compare our RM with FFM and with a combination of both models.

MATERIALS AND METHODS

Study Population

A total of nine patients (median age 11.5 years) with CP and 12 age-/height-/weight-matched healthy controls (median age 11.1 years) were prospectively enrolled between 2018 and 2020 after ethics committee approval (**Table 1**). All patients were into consideration for corrective surgery. They had fixed contractures with a functional component contributing to equinus gait. Six patients were diagnosed with unilateral spastic hemiparesis and 3 with spastic diparesis. At the time of the study, the MAS ranged between 0 and 2; only two hemiparetic patients had MAS of 0. The Gross Motor Function Classification System (GMFCS) level was mostly I, only two patients had level II, and another had level III motor function impediment. Passive range of motion (ROM) and manual muscle testing (MMT) of the knee joint was reported in **Table 2**. Informed consent was obtained from the parents of the participants and, additionally, from the 12 years old or older participants at the time of examination. Exclusion criteria were a history of surgery on the affected limb(s), claustrophobia, and failing to follow instructions during the acquisition.

Image Acquisition

The same scanner and the same acquisition parameters were used throughout the study. MRI exams were performed with a 3T whole-body scanner (Siemens Prisma, Siemens Healthineers, Erlangen, Germany). The patients and healthy controls were

TABLE 2 | Patient characteristics, results of the clinical examination, passive range of motion, manual muscle testing, and MRI fat fraction.

Patient	Patient Characteristics				Spasticity ‡			PROM		MMT^		Fat Fraction*		
	BMI percentile†	CP	GMFCS	More affected side	PF (at 90° KF)	PF (at KE)	KF	DF (at KE)	KE (at HE)	PF	KF	S	GM	GL
1	5	Unilateral	I	Right	1+	1+	0	−10°	0°	2+	5	6.8	5.5	16.7
2	95	Unilateral	I	Right	0	1	0	−20°	−15°	2+	4	24.4	16.5	16.5
3	23	Unilateral	I	Right	1+	1	0	15°	0°	2+	5	7.1	7.7	19.2
4	49	Bilateral	I	Left	1	1	0	−10°	−20°	2+	4	9.4	8.1	5.9
5	3	Bilateral	II	Left	1+	4	2	−30°	−5°	3	4	13.6	13.5	11.7
6	4	Bilateral	III	Left	1	1	1	−5°	−10°	3+	3	6	8.1	7.4
7	71	Unilateral	II	Left	2	2	0	10°	0°	2	4	5.1	5	4.7
8	61	Unilateral	I	Right	1	1	0	0°	10°	2+	5	4.1	3.6	4.3
9	5	Unilateral	I	Left	0	0	0	10°	5°	2+	4	3.3	8.2	5.1

PROM, passive range of motion; MMT, manual muscle testing; BMI, body mass index; CP, cerebral palsy; GMFCS, Gross Motor Function Classification System; PF, plantarflexor muscles; KF, knee flexion; KE, knee extension; DF, dorsiflexion; HE, hip extension; S, Soleus; GM, Gastrocnemius medialis; GL, Gastrocnemius lateralis.

† BMI percentile: underweight <5; 5 ≤ normal weight < 85; 85≤ overweight < 95; 95 ≤ obesity.

‡ Modified Ashworth Scale.

^ Medical Research Council scale.

**Fat fraction percentage in calf muscles of the affected legs of unilateral and more affected legs of the bilateral CP patients.*

positioned supine on the patient table, and the lower extremity was restrained with straps at a comfortable resting angle. A Siemens 18-element-body array coil was placed on the lower leg. A three-dimensional (3D) multi-echo gradient-echo (Dixon) sequence was used to reconstruct fat-only and water-only images: 6 echoes: echo times (TEs) 1.41/2.87/4.33/5.79/7.25/8.71 ms, repetition time (TR) = 20 ms, voxel size 1.1 × 1.1 × 3.0 mm^3, reconstructed matrix = 320 × 190 × 96, flip angle = 12°, acceleration factor 2, acquisition time 4 min 49 s (30). Images were acquired without contrast material and without anesthesia or sedation. Children were offered the possibility of visual or audio entertainment during the examination to improve compliance, and in case of suboptimal image quality, the corresponding acquisition was repeated.

Image Segmentation and Analysis

The images of the second echo (TE = 2.87 ms) from the multi-echo gradient-echo Dixon acquisition were used for the muscle segmentation. An in-house developed segmentation tool was employed for manual contour delineation. An experienced radiologist (T.A.D.) was responsible for all segmentations. The entire volumes of the calf muscles—soleus (S), gastrocnemius medialis (GM), and gastrocnemius lateralis (GL)—were segmented for both legs (**Figure 1**). Segmentations were then reviewed by one of the authors (C.W.) to check for errors.

All ROI margins were eroded by one voxel to reduce partial volume effects from adjacent adipose tissue and to prevent possible inadvertent overlaps between ROIs.

Water-only images, fat-only images, and fat fraction maps (defined as the signal intensity of the fat-only images divided by the sum of the signal intensities of fat-only and water-only images) were calculated online by the scanner software for all calf muscles bilaterally (30) (**Figure 1**).

Texture analysis was applied to the second echo images of the Dixon image series (TE = 2.87 ms). To improve the reproducibility and robustness of radiomics features, the voxel intensity range was normalized and quantized to 128 gray levels (31–33).

Radiomics features were extracted using Python version 3.8 (www.python.org) and the *PyRadiomics* package version 3.0 (34). A total of 107 features were extracted for each ROI.

Feature Selection and Model Building

Feature selection and dimension reduction methods were applied to prevent overfitting (35). Pearsons's correlation coefficient was used to test collinearity, and a heatmap was generated to demonstrate the collinearity between all extracted features. The least absolute shrinkage and selection operator (LASSO) regression and 10-fold cross-validation were employed to reduce the high dimension of all extracted features and select the most robust prognostic features among them as recommended (36). Bayesian information criterion was used for final feature selection. *The Image Biomarker Standardization Initiative (IBSI)* reference manual was used for feature definitions and calculations (37).

Logistic regression was used to build prediction models to predict the CP-affected muscle, i.e., spasticity. The radiomics model (RM), the fat fraction model (FFM), and a combination of both models (CM) were built for each one of the calf muscles separately. The receiver operating characteristic (ROC) curve and their respective area under the curve (AUC) values were used to evaluate model performance. A DeLong's test was used to compare performances of the three models (i.e., AUC values) within a muscle. A goodness-of-fit test was employed to assess how well the models were fitted. To compare the agreement between the actual and the predicted outcome, calibration curves were generated for the final models.

The feature selection and model building steps were performed in the 12 dominant legs of volunteers, 6 affected legs

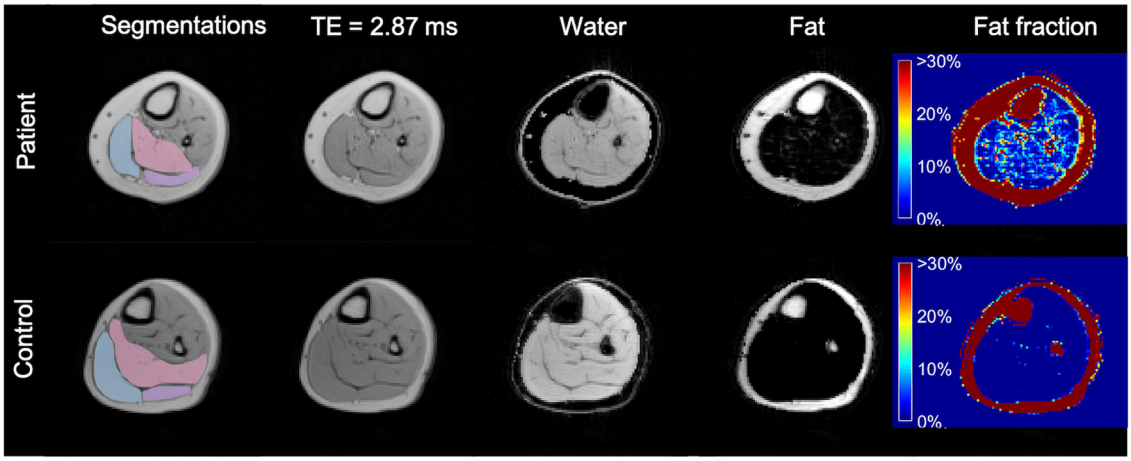


FIGURE 1 | Axial MR images of the more affected calf of a patient (diparetic, boy, 11 years) and the dominant calf of healthy control (a typically developing 11 years boy with similar BMI). First column: segmented ROIs for soleus (pink), gastrocnemius medialis (light blue), and gastrocnemius lateralis (lilac); second column: 2nd echo image from the Dixon data set; third column: water-only image calculated from the Dixon dataset; fourth column: fat-only image calculated from the Dixon dataset; fifth column: fat fraction map ranging from 0 to 100% calculated from the Dixon dataset, showing a higher fat fraction in the CP patient. TE, Echo Time.

of patients with unilateral spastic hemiparesis, and the 3 more affected legs of bilaterally affected patients.

Statistical Analysis

Statistical analysis was performed with Stata/IC 15.1 (StataCorp LP, College Station, Texas), and the *lassopack* (38) package was used. All continuous data, i.e., age, height, BMI, were given as either means and standard deviation or median and interquartile range. The group differences (CP vs. healthy) were assessed using Student's *t*-test or chi-squared test, where appropriate. Shapiro-Wilk test was used to assess the normality of the distributions. Alpha level was set to 0.05.

RESULTS

Study Population and Fat Fraction

In total, 21 lower limbs (9 affected, 12 healthy) were included.

Mean FF values were higher in the affected legs of the patients with unilateral spastic hemiparesis and more affected legs of patients with spastic diparesis than dominant legs of the healthy controls. While CP-affected and more affected legs had FF_S 9.1%, FF_{GM} 8.5% and FF_{GL} 10.2%, the control legs had FF_S 3.3%, FF_{GM} 4.1%, FF_{GL} 6.6% (*p*-values 0.009, 0.005, 0.116, respectively). The FF difference between dominant and non-dominant legs in the control group was not apparent for S and GM muscles, whereas the difference was pronounced for GL. Similarly, the FF values in the contralateral leg of patients were lower than affected/more affected leg for S and GM muscles, but this difference was not evident for GL. An overview of the FF values is given in **Figure 2**.

Feature Selection

Feature selection was performed for the entire dataset. All selected features and their values are reported in **Supplementary Tables 1–3**. The LASSO regression model successfully reduced the dimensionality of 107 features and selected the most robust ones (**Supplementary Figure 1**). Since LASSO also accounts for collinearity, no further steps were taken in the Pearson correlation coefficient. A correlation heatmap of all extracted features shown in **Figure 3** and depicts little redundancy. All chosen texture features with a non-zero coefficient in the LASSO regression are reported in **Table 3**. A different set of features were selected for each calf muscle. The selected features belonged to 2D shape-based (maximum 2D diameter row, surface volume ratio), gray level co-occurrence matrix (information correlation 1, cluster shade), and gray level size zone matrix (small area low gray-level emphasis, small area emphasis) feature classes (**Table 3** and **Supplementary Data**).

Predictive Models

An RM, an FFM, and a CM were built for each muscle separately. All model performances are reported in **Figure 4**. Based on the ROC analysis, the performance of RM was excellent for soleus and gastrocnemius medialis and very good for gastrocnemius lateralis (AUC_S 0.92; AUC_{GM} 0.92, AUC_{GL} 0.82). The FFM always showed good performance for soleus and gastrocnemius medialis and moderate performance for gastrocnemius lateralis (AUC_S 0.88; AUC_{GM} 0.85; AUC_{GL} 0.69). The combination of both models always had a better performance than RM according to ROC analysis (AUC_S 0.95; AUC_{GM} 0.93; AUC_{GL} 0.83) (**Figure 5**). The sensitivity of RM was between 67 and 89%, and specificity was between 83 and 100%. The accuracy of RM was always higher than FFM, and the CM model accuracy was either better than or comparable to RM (**Figure 4**). The calibration curves of the final combination models showed a high level of agreement between the prediction of the affected muscle and actually affected muscle (**Supplementary Figure 2**).

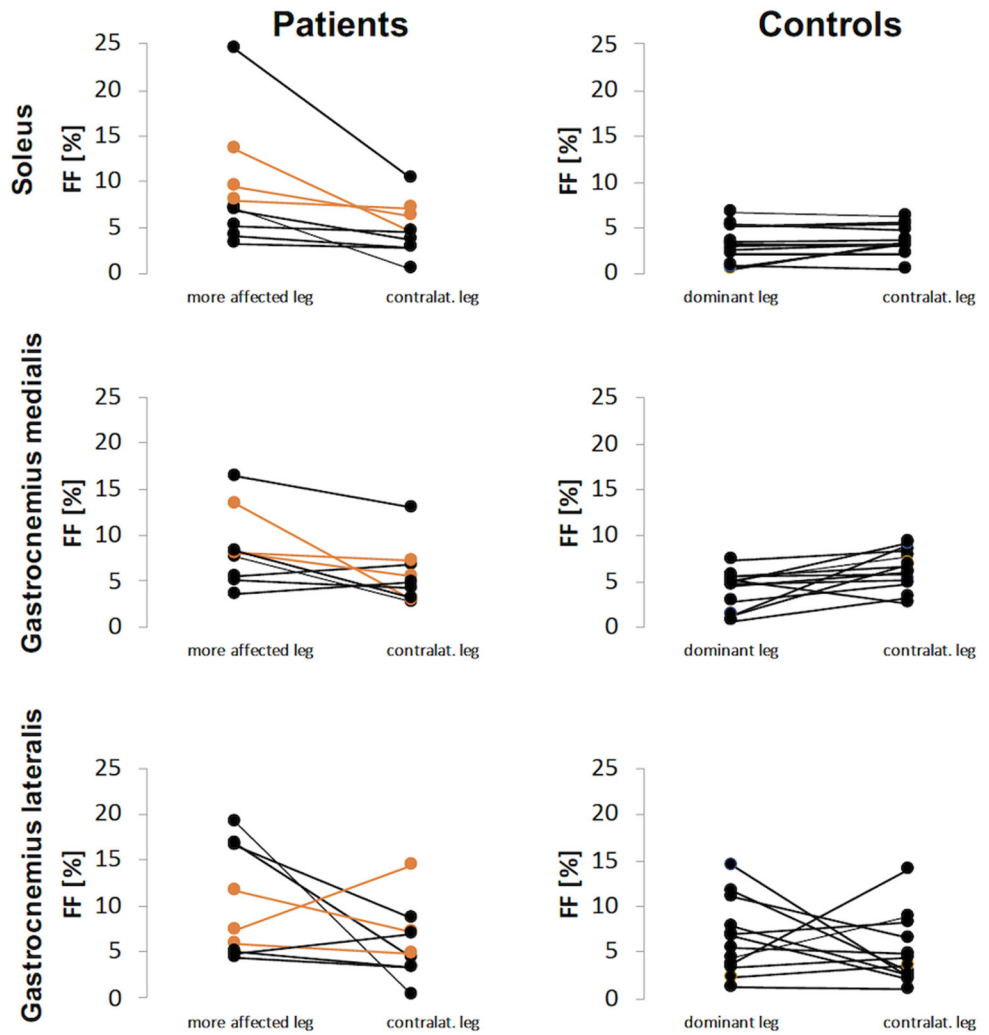


FIGURE 2 | Fat fraction (FF) values for patients and controls for each calf muscle. The FF results of patients with diparesis are depicted in orange.

DISCUSSION

In this study, we employed radiomics analysis of MR images in CP patients and healthy controls to discriminate affected muscles from healthy ones. We compared the performance of the RM with the FFM as well as with the combination of both. Our radiomics analysis yielded a better performing model than the FF analysis. Moreover, we found that the combination of both models always performed better.

So far, quantitative MRI techniques—in particular, FF analysis—have been employed to explore disease severity or to monitor treatment response in muscle dystrophies (5–8). Muscle FF analysis has been rarely used to assess the functional capacity of CP patients, and researchers have reported higher fat quantity in muscles of CP patients than in a healthy population (13, 14). In our analysis, the patient who had the highest FF results for all muscles was obese with a BMI of 95 percentile. On the other hand, the patient with the second highest FF results was underweight with BMI 3 percentile and had the most severe fixed contracture. All other patients were normal weighted and had lower FF values than aforementioned 2 patients. In line with previous studies, our FF results were higher in the patient group than the control group regardless of their BMI. The FF values in our patient group are similar to the values in hemiparetic children reported by D'Souza et al. (14), and lower than the values in biparetic young adults that reported by Noble et al. (13). Although FF could be a useful tool to assess muscle diseases, it is usually restricted to a single value estimation of each muscle, disregarding the inhomogeneity of muscle structure, especially in the presence of pathology.

A less expensive and easily accessible alternative to MRI is ultrasonography (US). The US can be employed to detect basic structural muscle changes and assess muscle volume in CP patients (39, 40). Nonetheless, it is highly user-dependent and does not provide a global view of all muscles (39). Although new emerging US techniques, e.g., sheer-wave elastography, can

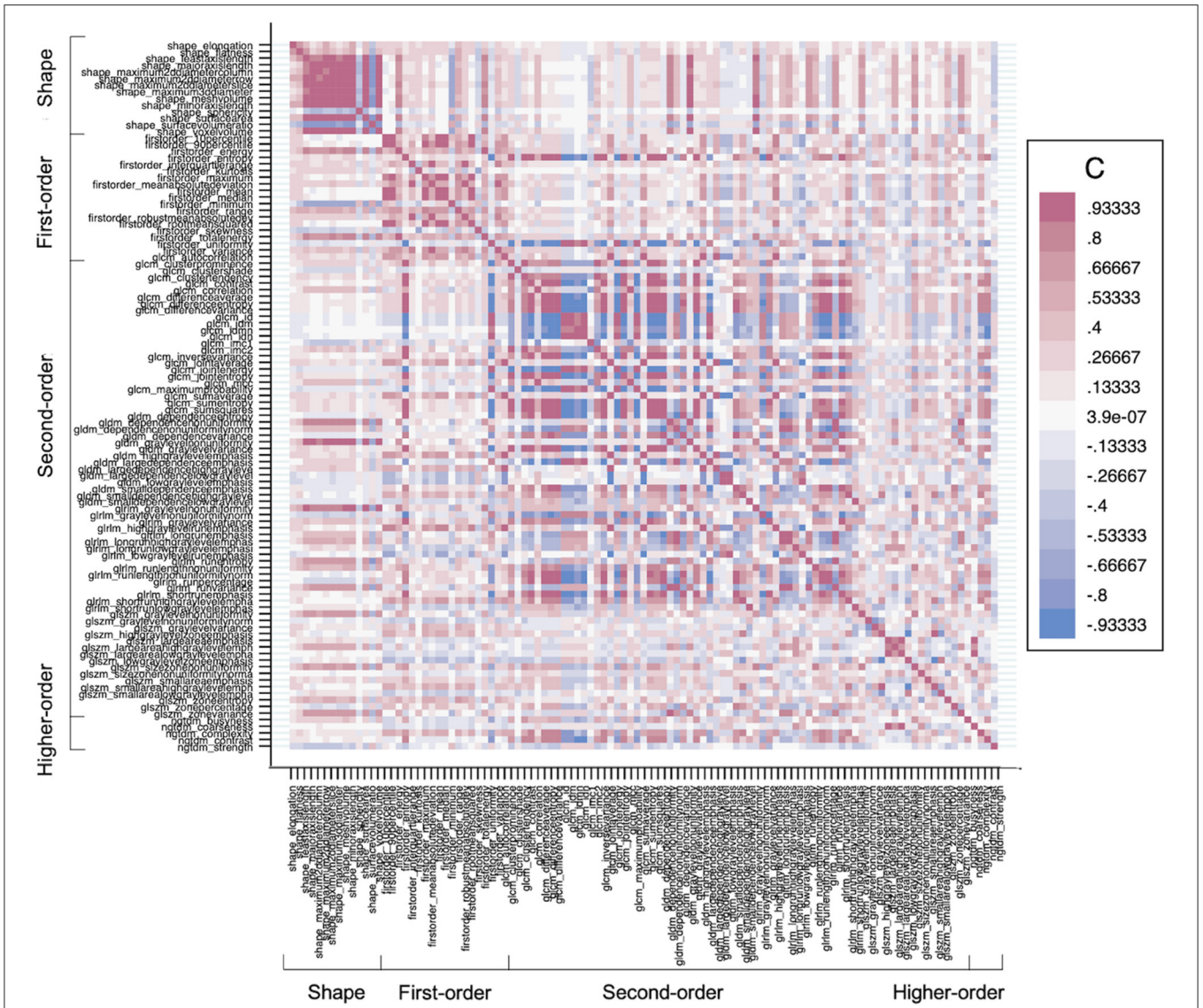


FIGURE 3 | A heatmap demonstrates the collinearity between all extracted features—correlation coefficient (C) range between −1 and +1. Red depicts the perfect positive correlation, and blue depicts the perfect negative correlation, and all other colors depict correlation in between on the heat map. The higher the C in each direction, the more redundant the feature is.

make functional predictions (40), the composition of muscle, in particular the fat fraction, cannot be detected using only US (39). On the other hand, MRI can be used not only for global assessment of muscles but also for compositional assessment.

Texture analysis of diagnostic images is a non-invasive tool that can shed light on the underlying pathophysiology. Radiomics analysis can guide biopsies and play a role in following up the disease progression by longitudinal radiomics analysis, so called delta radiomics (41). MRI texture analysis has demonstrated to be a potential tool to evaluate neuromuscular muscle disorders in animal models (21, 22, 25–27), and preliminary studies already established some texture biomarkers for assessing disease progression in a dog model of muscular dystrophies (25, 27). Moreover, recent studies revealed that MRI texture analysis could help investigate the effects of repetitive forces in healthy athletes by detecting texture changes due to muscle hypertrophy (23, 24). So far, only a few studies have applied texture analysis to various pathologic skeletal muscles, e.g., muscle dystrophy, in human subjects (28, 29). Researchers explored the correlations between texture analysis and the disease status and found muscle texture features helpful for objective evaluation of MRI (28, 29). To our knowledge, our study is the first one that applied radiomics analysis to the MRI of skeletal muscles of children with CP.

In our study, we applied an FF analysis along with texture analysis to predict CP affected muscles. The FF analysis has

TABLE 3 | All selected texture features and their values, LASSO coefficients, and IBSI reference values.

Muscle	Selected Features	Feature Value		LASSO coefficient	IBSI Reference value*	Student's t-test *P* value
		Healthy ($n = 12$)	Cerebral palsy ($n = 9$)			
Soleus	GLCM cluster shade	−0.1 ± 0.3	−0.1 ± 0.3	−0.17	7.0	0.560
	GLCM information correlation 1	−0.3 ± 0.1	−0.2 ± 0.1	0.12	−0.1	0.043
	GLSZM Small area low gray-level emphasis	0.4 ± 0.1	0.2 ± 0.2	−0.62	0.02	0.044
Gastrocnemius medialis	Shape maximum 2D diameter row	52.9 ± 7.5	40.1 ± 7.4	−0.01	13.1	< 0.001
	GLSZM small area emphasis	0.6 ± 0.1	0.5 ± 0.1	−0.07	0.3	0.017
Gastrocnemius lateralis	Shape surface volume ratio	0.4 ± 0.1	0.5 ± 0.1	0.60	0.7	0.005

Data are means ± standard deviations.
LASSO, least absolute shrinkage and selection operator; IBSI, Image Biomarker Standardization Initiative; GLCM, gray level co-occurrence matrix; GLSZM, gray level size zone matrix.
**Reference values that are reported in IBSI reference manual for digital phantom at the highest consensus level.*

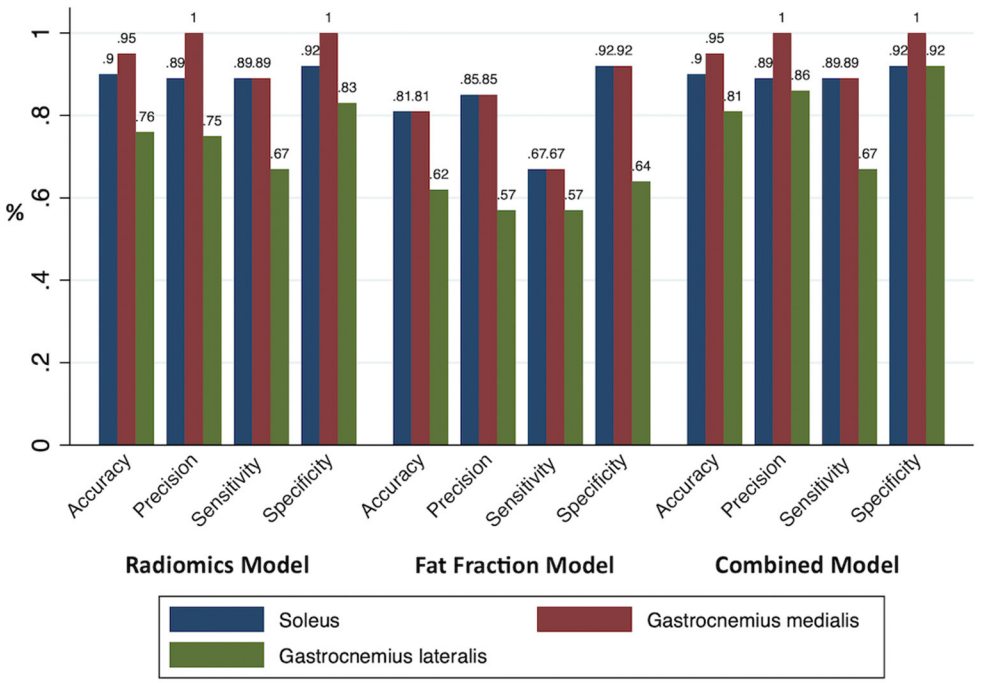


FIGURE 4 | All model performances. The accuracy, precision, sensitivity, and specificity of the prediction models were based on the radiomics, fat fraction, and combined model built for each muscle.

been performed based on multi-echo Dixon acquisition with 6 echoes, as at least 3 echoes (Dixon points) are recommended to overcome main field inhomogeneities (42). Images of the second echo (TE = 2.87 ms) were used for feature extraction. The RM always performed better than FFM in discriminating affected muscles with an accuracy between 76 and 95%. Furthermore, the combination of those methods showed an excellent performance level with an accuracy between 81 and 95%. Although fat quantification with FF has a relevant role in evaluating the muscles of the CP patients, our FFM resulted in only moderate/good performance level with accuracy between 62 and 81%. This might be due to the calf muscles' immediate proximity to subcutaneous fat tissue, which was especially prominent on GL FF analysis. Despite this issue being addressed by eroding the ROIs by one voxel, the performance of the FFM was lower than RM or CM. On the other hand, RM always reached a high-performance level. Therefore, combining radiomics and FF methods might especially be recommended for assessing muscles adjacent to subcutaneous fat tissue.

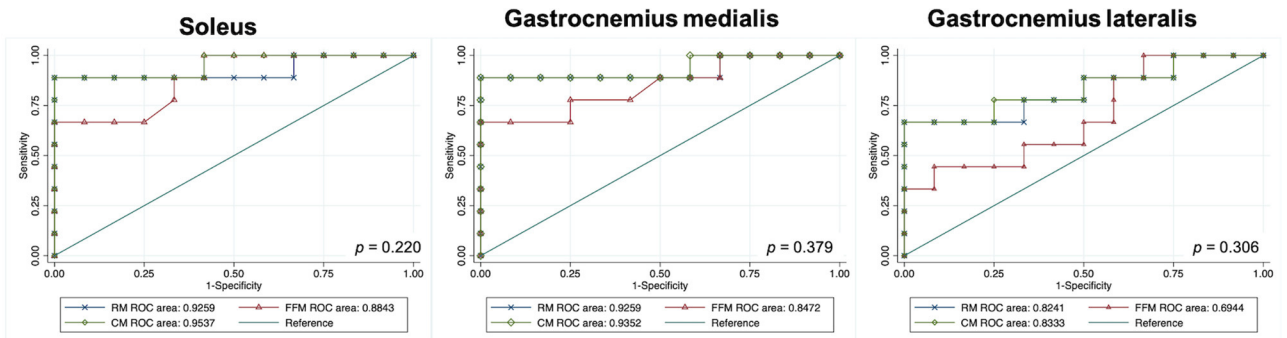


FIGURE 5 | Graphs show receiver operating characteristic curves indicating the accuracy of models for predicting disease in children with cerebral palsy. DeLong's test *P* values. GM, gastrocnemius medialis; GL, gastrocnemius lateralis; RM, radiomics model; FFM, fat fraction model; CM, combined model; RM, radiomics model; FFM, fat fraction model; CM, combined model; ROC, receiver operating characteristic.

We employed the LASSO regression for dimension reduction and feature selection since those steps are the pillars of the texture analysis and, consequently, help avoid overfitting (35, 36). Higher-order statistics features were eliminated by LASSO, presumably due to their sensitivity to noise. In contrast, second-order statistics and the shape features are less affected by noise and, therefore, more robust (20, 37). Hence, our LASSO analysis mostly selected the shape and second-order features instead of higher-order statistics features. Shape features define the two-dimensional size and shape of the ROI (37). These features are independent of the gray level intensity distribution in the ROI (37). Shape features were the one of the most successful features in our analysis, and they could be employed as an imaging biomarker for CP patients since the normal shape of the muscle can be drastically altered due to spasticity and contractions (4). The co-occurrence matrix depicts the frequency of a pair of pixels with the same value in a specified spatial range within an ROI (37). Co-occurrence matrix features were also successful in our study since they can reveal the texture heterogeneity due to fat infiltration within muscle tissue in CP patients. Gray level distance zone-based features depict the frequency of groups (zones) of the same gray-level appear in every direction within a voxel (37). Other successful features were belonging to this group. Fat infiltration can change the gray levels, and since this feature takes into account the neighboring voxel relations, it can reveal the extension of the fat infiltration. The RM reached a high level of performance to discriminate CP-affected muscles from the normal ones. It is of particular interest that although the spasticity was reduced at the time of the imaging, the muscle alterations, which could be due to remaining contractions or subtle structural changes, were successfully detected in radiomics analysis.

Our study had some limitations. Firstly, our study population was small. However, acquiring MRI data from a specific cohort, i.e., children with spasticity who can comply with MRI examination without sedation, was particularly challenging. Secondly, a single observer performed all segmentations, and segmentations were done manually. Yet, another observer controlled the segmentations against errors. It is well-known that inter-rater agreement is low in segmentation tasks, and although automated segmentation methods are desirable, manual segmentation by a single reader still provides a high degree of reliability for the reproducibility of radiomics features (20). Our patient group had fixed calf contractures, and it is known that muscle changes in patients with fixed contractures are more dramatic than the patients with dynamic contractures (2). Therefore, studies in a heterogeneous patient group needed to demonstrate the applicability of our model on less pronounced contractures. Moreover, we did not have a radiology-pathology correlation; therefore, the true relation between radiomics or fat fraction analysis and muscle histopathology still needs to be elucidated. Nevertheless, radiomics analysis of skeletal muscles is a promising tool to provide non-invasive tissue characterization and reduce muscle biopsies since it is particularly important to avoid unnecessary interventions in the pediatric population, and muscle biopsies usually fail to capture tissue heterogeneity or to reflect the entire tissue. Still, radiology-pathology correlation studies are required to understand the relationship between muscle histopathology and imaging biomarkers. Lastly, we have not tested our model on an independent external dataset. Nevertheless, we did internal validation with 10-fold cross-validation; further studies are required for external validation of the proposed model.

In conclusion, the combination of MRI quantitative fat fraction analysis and texture analysis of muscles in CP patients is a promising tool to evaluate skeletal muscle involvement of the disease in a non-invasive manner. In the long term, our model could be integrated into clinical decision-making systems, and a similar approach might be used to assess other muscle diseases. Further investigations in a large cohort of patients with CP are needed to optimize and validate our proposed model in a clinical setting.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because restrictions apply to DICOM data originating from the

MRI scanner. Requests to access the datasets should be directed to CW, claudia.weidensteiner@unibas.ch.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission Nordwest-und Zentralschweiz, Basel, Switzerland. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

TA, XD, FS, and CW: study concept, study design and data acquisition, analysis, and interpretation. RB and ER: clinical studies. TA, XD, FS, CW, and MG: diagnostic imaging. FS: segmentation tool. TA and XD: statistical analysis. TA, XD, and CW: literature search. TA: first draft of the manuscript. All authors: manuscript revision for important intellectual content and manuscript final version approval. All authors agree to ensure any questions related to the work are appropriately resolved.

FUNDING

This study was supported by the Swiss National Science Foundation (Grant No. 173292).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.633808/full#supplementary-material

REFERENCES

1. Morris C. Definition and classification of cerebral palsy: a historical perspective. *Dev Med Child Neurol*. (2007) 49(Suppl. 2):3–7. doi: 10.1111/j.1469-8749.2007.tb12609.x
2. Graham HK, Rosenbaum P, Paneth N, Dan B, Lin JP, Damiano DiL, et al. Cerebral palsy. *Nat Rev Dis Prim*. (2016) 2:15082. doi: 10.1038/nrdp.2016.5
3. Pons C, Rémy-Néris O, Médée B, Brochard S. Validity and reliability of radiological methods to assess proximal hip geometry in children with cerebral palsy: a systematic review. *Dev Med Child Neurol*. (2013) 55:1089–102. doi: 10.1111/dmcn.12169
4. Oberhofer K, Stott NS, Mithraratne K, Anderson IA. Subject-specific modelling of lower limb muscles in children with cerebral palsy. *Clin Biomech*. (2010) 25:88–94. doi: 10.1016/j.clinbiomech.2009.09.007
5. Leung DG. Advancements in magnetic resonance imaging-based biomarkers for muscular dystrophy. *Muscle and Nerve*. (2019) 60:347–60. doi: 10.1002/mus.26497
6. Bray TJP, Chouhan MD, Punwani S, Bridge A, Hall-Craggs MA. Fat fraction mapping using magnetic resonance imaging: Insight into pathophysiology. *Br J Radiol*. (2018) 91:20170344. doi: 10.1259/bjr.20170344
7. Nuñez-Peralta C, Alonso-Pérez J, Díaz-Manera J. The increasing role of muscle MRI to monitor changes over time in untreated and treated muscle diseases. *Curr Opin Neurol*. (2020) 33:611–20. doi: 10.1097/WCO.0000000000000851
8. Burakiewicz J, Sinclair CDJ, Fischer D, Walter GA, Kan HE, Hollingsworth KG. Quantifying fat replacement of muscle by quantitative MRI in muscular dystrophy. *J Neurol*. (2017) 264:2053–67. doi: 10.1007/s00415-017-8547-3
9. Johnston JH, Kim HK, Merrow AC, Laor T, Serai S, Horn PS, et al. Quantitative skeletal muscle MRI: part 1, derived T2 fat map in differentiation between boys with duchenne muscular dystrophy and healthy boys. *Am J Roentgenol*. (2015) 205:W207–15. doi: 10.2214/AJR.14.13754
10. Kim HK, Serai S, Lindquist D, Merrow AC, Horn PS, Kim DH, et al. Quantitative skeletal muscle MRI: part 2, MR spectroscopy and T2 relaxation time mapping- comparison between boys with duchenne muscular dystrophy and healthy boys. *Am J Roentgenol*. (2015) 205:W216–23. doi: 10.2214/AJR.14.13755
11. Ahlberg G, Jakobsson F, Fransson A, Mority A, Borg K, Edstrom L. Distribution of muscle degeneration in Welander distal myopathy—A magnetic resonance imaging and muscle biopsy study. *Neuromuscul Disord*. (1994) 4:55–62. doi: 10.1016/0960-8966(94)90048-5
12. Klonowski W, Kuraszkiewicz B, Kaminska A, Kostera-Pruszczyk A. Texture analysis of dystrophic muscle biopsies. *bioRxiv*. (2018) 273771. doi: 10.1101/273771
13. Noble JJ, Charles-Edwards GD, Keevil SF, Lewis AP, Gough M, Shortland AP. Intramuscular fat in ambulant young adults with bilateral spastic cerebral palsy. *BMC Musculoskelet Disord*. (2014) 15:1–8. doi: 10.1186/1471-2474-15-236
14. D'Souza A, Bolsterlee B, Lancaster A, Herbert R. Intramuscular fat in children with unilateral cerebral palsy. *Clin Biomech*. (2020) 80:105183. doi: 10.1016/j.clinbiomech.2020.105183
15. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. (2016) 278:563–77. doi: 10.1148/radiol.2015151169
16. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. (2014) 5:4644. doi: 10.1038/ncomms5644
17. Beig N, Khorrami M, Alilou M, Prasanna P, Braman N, Orooji M, et al. Perinodular and intranodular radiomic features on lung CT images distinguish adenocarcinomas from granulomas. *Radiology*. (2018) 290:783–92. doi: 10.1148/radiol.2018180910
18. Akinci D'Antonoli T, Farchione A, Lenkowicz J, Chiappetta M, Cicchetti G, Martino A, et al. CT radiomics signature of tumor and peritumoral lung parenchyma to predict nonsmall cell lung cancer postsurgical recurrence risk. *Acad Radiol*. (2020) 27:497–507. doi: 10.1016/j.acra.2019.05.019
19. De Certaines JD, Larcher T, Duda D, Azzabou N, Eliat P-A, Escudero LM, et al. Application of texture analysis to muscle MRI: 1-What kind of information should be expected from texture analysis? *EPJ Nonlinear Biomed Phys*. (2015) 3:1–14. doi: 10.1140/epjnbp/s40366-015-0017-1
20. Lerski RA, de Certaines JD, Duda D, Klonowski W, Yang G, Coatrieux JL, et al. Application of texture analysis to muscle MRI: 2 – technical recommendations. *EPJ Nonlinear Biomed Phys*. (2015) 3:2. doi: 10.1140/epjnbp/s40366-015-0018-0
21. Martins-Bach AB, Malheiros J, Matot B, Martins PCM, Almeida CF, Caldeira W, et al. Quantitative T2 combined with texture analysis of nuclear magnetic resonance images identify different degrees of muscle involvement in three mouse models of muscle dystrophy: Mdx, Largemydand mdx/Largemyd. *PLoS ONE*. (2015) 10:117835. doi: 10.1371/journal.pone.0117835
22. Mahmoud-Ghoneim D, Cherel Y, Lemaire L, de Certaines JD, Maniere A. Texture analysis of magnetic resonance images of rat muscles during atrophy and regeneration. *Magn Reson Imaging*. (2006) 24:167–71. doi: 10.1016/j.mri.2005.10.002
23. Nketiah G, Savio S, Dastidar P, Nikander R, Eskola H, Sievänen H. Detection of exercise load-associated differences in hip muscles by texture analysis. *Scand J Med Sci Sport*. (2015) 25:428–34. doi: 10.1111/sms.12247
24. Sikiö M, Harrison LCV, Nikander R, Ryymin P, Dastidar P, Eskola HJ, et al. Influence of exercise loading on magnetic resonance image

texture of thigh soft tissues. *Clin Physiol Funct Imaging*. (2014) 34:370–6. doi: 10.1111/cpf.12107
25. Wang J, Fan Z, Vandenborne K, Walter G, Shiloh-Malawsky Y, An H, et al. A computerized MRI biomarker quantification scheme for a canine model of Duchenne muscular dystrophy. *Int J Comput Assist Radiol Surg*. (2013) 8:763–74. doi: 10.1007/s11548-012-0810-6
26. Pratt SJP, Xu S, Mullins RJ, Lovering RM. Temporal changes in magnetic resonance imaging in the mdx mouse. *BMC Res Notes*. (2013) 6:3–7. doi: 10.1186/1756-0500-6-262
27. Fan Z, Wang J, Ahn M, Shiloh-Malawsky Y, Chahin N, Elmore S, et al. Characteristics of magnetic resonance imaging biomarkers in a natural history study of golden retriever muscular dystrophy. *Neuromuscul Disord*. (2014) 24:178–91. doi: 10.1016/j.nmd.2013.10.005
28. Škoch A, Jirák D, Vyhnanovská P, Dezortová M, Fendrych P, Rolencová E, et al. Classification of calf muscle MR images by texture analysis. *Magn Reson Mater Phys Biol Med*. (2004) 16:259–67. doi: 10.1007/s10334-004-0032-1
29. Herlidou S, Rolland Y, Bansard JY, Le Rumeur E, de Certaines JD. Comparison of automated and visual texture analysis in MRI: characterization of normal and diseased skeletal muscle. *Magn Reson Imaging*. (1999) 17:1393–7. doi: 10.1016/S0730-725X(99)00066-1
30. Henninger B, Zoller H, Kannengiesser S, Zhong X, Jaschke W, Kremser C. 3D Multiecho dixon for the evaluation of hepatic iron and fat in a clinical setting. *J Magn Reson Imaging*. (2017) 46:793–800. doi: 10.1002/jmri.25630
31. Kociołek M, Strzelecki M, Obuchowicz R. Does image normalization and intensity resolution impact texture classification? *Comput Med Imaging Graph*. (2020) 81:101716. doi: 10.1016/j.compmedimag.2020.101716
32. Collewet G, Strzelecki M, Mariette F. Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. *Magn Reson Imaging*. (2004) 22:81–91. doi: 10.1016/j.mri.2003.09.001
33. Li Q, Bai H, Chen Y, Sun Q, Liu L, Zhou S, et al. A Fully-automatic multiparametric radiomics model: towards reproducible and prognostic imaging signature for prediction of overall survival in glioblastoma multiforme. *Sci Rep*. (2017) 7:14331. doi: 10.1038/s41598-017-14753-7
34. Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res*. (2017) 77:e104–7. doi: 10.1158/0008-5472.CAN-17-0339
35. Harrell FE. *Regression Modeling Strategies*. New York, NY: Springer-Verlag (2015).
36. Zwanenburg A. Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis. *Eur J Nucl Med Mol Imaging*. (2019) 46:2638–55. doi: 10.1007/s00259-019-04391-8
37. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology*. (2020) 295:328–38. doi: 10.1148/radiol.2020191145
38. Ahrens A, Hansen CB, Schaffer ME. lassopack: Model selection and prediction with regularized regression in Stata. *Stata J Promot Commun Stat Stata*. (2020) 20:176–235. doi: 10.1177/1536867X20909697
39. Mathewson MA, Lieber RL. Pathophysiology of muscle contractures in cerebral palsy. *Phys Med Rehabil Clin N Am*. (2015) 26:57–67. doi: 10.1016/j.pmr.2014.09.005
40. van Holsbeeck M, Soliman S, Van Kerkhove F, Craig J. Advanced musculoskeletal ultrasound techniques: what are the applications? *Am J Roentgenol*. (2021) 216:436–45. doi: 10.2214/AJR.20.22840
41. Tomaszewski MR, Gillies RJ. The biological meaning of radiomic features. *Radiology*. (2021) 2021:202553. doi: 10.1148/radiol.2021202553
42. Strijkers GJ, Araujo ECA, Azzabou N, Bendahan D, Blamire A, Burakiewicz J, et al. Exploration of new contrasts, targets, and MR imaging and spectroscopy techniques for neuromuscular disease – a workshop report of working group 3 of the biomedicine and molecular biosciences COST action BM1304 MYO-MRI. *J Neuromuscul Dis*. (2019) 6:1–30. doi: 10.3233/JND-180333

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

REVIEW
published: 25 March 2021
doi: 10.3389/fneur.2021.625308

Overview of MR Image Segmentation Strategies in Neuromuscular Disorders

Augustin C. Ogier[1]*, *Marc-Adrien Hostin*[1,2], *Marc-Emmanuel Bellemare*[1] and *David Bendahan*[2]

[1] Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France, [2] Aix Marseille Univ, CNRS, CRMBM, UMR 7339, Marseille, France

Neuromuscular disorders are rare diseases for which few therapeutic strategies currently exist. Assessment of therapeutic strategies efficiency is limited by the lack of biomarkers sensitive to the slow progression of neuromuscular diseases (NMD). Magnetic resonance imaging (MRI) has emerged as a tool of choice for the development of qualitative scores for the study of NMD. The recent emergence of quantitative MRI has enabled to provide quantitative biomarkers more sensitive to the evaluation of pathological changes in muscle tissue. However, in order to extract these biomarkers from specific regions of interest, muscle segmentation is mandatory. The time-consuming aspect of manual segmentation has limited the evaluation of these biomarkers on large cohorts. In recent years, several methods have been proposed to make the segmentation step automatic or semi-automatic. The purpose of this study was to review these methods and discuss their reliability, reproducibility, and limitations in the context of NMD. A particular attention has been paid to recent deep learning methods, as they have emerged as an effective method of image segmentation in many other clinical contexts.

Keywords: MRI, neuromuscular disorders, muscle, image segmentation, deep learning

OPEN ACCESS

Edited by:
Francesco Santini, University of Basel, Switzerland

Reviewed by:
Alberto Stefano Tagliafico, University of Genoa, Italy
Abramo Agosti, Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy

***Correspondence:**
Augustin C. Ogier
augustin.ogier@gmail.com

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: *02 November 2020*
Accepted: *08 February 2021*
Published: *25 March 2021*

Citation:
Ogier AC, Hostin M-A, Bellemare M-E and Bendahan D (2021) Overview of MR Image Segmentation Strategies in Neuromuscular Disorders. Front. Neurol. 12:625308. doi: 10.3389/fneur.2021.625308

1. INTRODUCTION

Neuromuscular pathologies are rare diseases that can occur in both children and adults. Very few therapeutic strategies have been proposed so far. Reliable outcome measures that could be sensitive enough to detect therapeutic effects are still missing. Diagnosis of neuromuscular pathologies is commonly based on clinical presentation, genetic testing, and histological assessment of muscle biopsy. Given its non-invasiveness and its ability to distinguish fat and muscle tissue, magnetic resonance imaging (MRI), and more particularly quantitative MRI (qMRI), has emerged in recent years as a tool of choice for the investigation of neuromuscular diseases (1). Over the last 20 years, research projects have been developed to define relevant and sensitive MRI parameters that could be used in the diagnostic classification and the follow-up of neuromuscular diseases (2–6). The initial approaches were based on a visual analysis of hypersignals illustrating pathological processes and on that basis alterations patterns have been proposed. In slowly progressive diseases, such as neuromuscular disorders, the sensitivity of such visual qualitative assessments is largely questionable and may not be powerful enough to identify mild changes in muscle function that occur from year to year.

More recently, qMRI have been used in order to generate parametric maps illustrating the various pathological processes occurring in skeletal muscle, i.e., mainly inflammation and fat infiltration (4, 5). Compared to visual scores, such a quantitative approach has paved the way of a more sensitive assessment of dystrophies and neuropathies. Beyond the diagnostic interest, these approaches provide sensitive and reproducible biomarkers, which have been used for follow-up studies (6–8). In addition to the generation of parametric maps, qMRI has to be combined to images segmentation if one intends to extract the relevant information within different regions of interest. Segmentation refers to the delineation of muscle regions of interest that must be distinguished from subcutaneous and perimuscular adipose tissues, on the one hand, and from bones, on the other hand. Segmentation in general and segmentation of MR images in particular is a time-consuming process so that automatic procedures are highly requested. However, automation of muscle segmentation in MR images is very challenging given the poor contrast between different muscles and the large variability of muscle shapes (9). In pathological situations, the challenge can be even higher given that borders between the different compartments can be hidden by a severe fat infiltration. Given the task complexity, most of the studies related to the investigation of neuromuscular diseases have been based on the manual segmentation of individual muscles or muscle compartments. One has to keep in mind that such an approach is operator dependent and time consuming. As a result, quantitative analyses have been mainly performed over a limited number of slices and not on the whole 3D datasets (4, 10) or on a limited number of individual muscles (6). In a few clinical studies in which the manual segmentation has been performed in the 3D field of view, an inconsistent distribution of MRI biomarker values has been interestingly reported along the proximo-distal axis (2, 11–13). In addition, it has also been documented that individual muscles could be affected differently and that this difference could also occur among patients and neuromuscular disorders (5, 14–16). These results clearly emphasized the need of reliable automatic 3D segmentation methods and the relevance of evaluating muscles individually, rather than by muscle groups.

Over the last 15 years, several automated methods have been reported in the literature with the aim of segmenting muscle groups or individual muscles in MR images. Although promising, most of these methods have been tested in MR images from healthy volunteers for which fat infiltration and atrophy were absent so that the corresponding confounding factors could not be taken into account. More recently, a few automatic methods have addressed the issue of segmenting MR images of patients with neuromuscular disorders but only for the delineation of muscle compartment.

Very recently, semi-automatic methods have been reported for individual muscles segmentations in order to reach an optimized balance between segmentation accuracy and user's dedication. These full 3D methods have been successfully used in a clinical context. A few limitations has to be acknowledged for these kinds of methods. They are still time consuming and require a manual initialization so that reliable full automatic segmentation methods are still warranted for individual muscles.

Deep learning methods have been very scarcely used in the field of neuromuscular disorders and considering the results obtained in other scientific fields, they might represent a very interesting alternative for a full-3D segmentation of MR images. One should keep in mind that large databases should be available for this kind of approach and this could be a limitation in rare diseases.

Manual segmentation methods are not applicable for 3D datasets and the follow-up of neuromuscular diseases. They have been recognized as time consuming (5 h per subject for the 3D manual segmentation of 4 muscles) and operator dependent (3.1% volume error for the *quadriceps femoris* in healthy subjects) (9). On that basis, these methods were beyond the scope of the present review. Considering that neuromuscular disorders have been mainly studied using MRI of thighs and legs, only the automated methods that have been used for the segmentation of lower limbs images are part of the scope of the present review.

The main aim of this review is to provide an overview of the methods dedicated to the segmentation of individual skeletal muscles on MR images and to discuss their validity and reliability. We pay a particular attention to the evolution of segmentation strategies, from the separation of muscle and fat deposits to the segmentation of individual muscles, together with the clinical potential and applicability in the context of neuromuscular disorders. Finally, we introduce insights into semi-automatic methods that could potentially break the barrier between research and clinics. These methods could provide clinicians with user-friendly tools that generate biomarkers for individual muscles over an entire 3D dataset. The emerging segmentation methods based on deep learning approaches have been included in a dedicated section as they are still emerging.

2. MUSCLE TISSUE SEGMENTATION ISSUES

2.1. Type of MR Images for Segmentation

Since the emergence of MRI, the quality and type of images available through this acquisition modality have greatly evolved and have consequently influenced automated segmentation methods. First segmentation methods had to deal with severe artifacts on MR images and hence segmentation of the contour of lower limbs and muscular region were a complicated task (17). Over the years, image quality has been dramatically improved with hardware and images techniques advancement. Intensity inhomogeneity correction method, such as N3 and more recently N4 algorithm (18), also strongly contributed to the improvement of image quality and such algorithms are now a common pre-processing step for muscle segmentation method. Segmentation methods for lower limbs MR images were first dedicated to T_1-weighted images, commonly used in clinics. The parametric images from qMRI sequences were then used as they may display different information regarding the nature of tissue (**Figure 3**). In most of the studies discussed in this review, parametric maps used for segmentation methods were extracted from multi-echo chemical shift-based water-fat separation MR sequences (19).

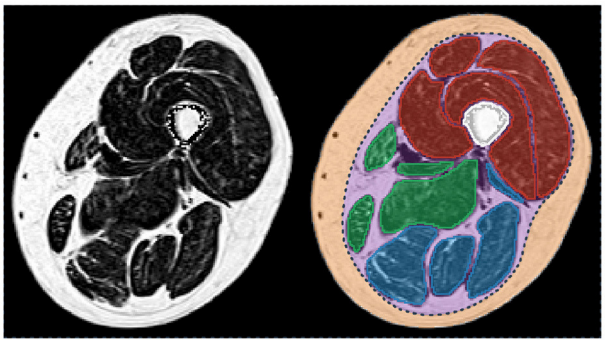


FIGURE 1 | Fat fraction map of a thigh of a patient with myotonic dystrophy type 1 **(left)** and corresponding segmentation of the principal regions of interest **(right)**. Femur is in white. The individual muscles of the knee extensors, knee flexors, and the medial compartment are in red, blue, and green, respectively. Subcutaneous (orange) and perimuscular (purple) tissues are separated by the *fascia lata* (dot line in dark blue). Patient participated in the randomized controlled trial OPTIMISTIC (5).

2.2. Regions of Interest

As illustrated in **Figure 1**, different tissues are visible in a MR image of a lower limb. For the sake of clarity, these tissues will be designated according to the nomenclature of (20). The subcutaneous adipose tissue (SAT) and the internal adipose tissue (IAT) are separated by the *fascia lata* for the thigh and by several deep fascias for the lower leg. Within the IAT, the intramuscular adipose tissue is defined as the adipose tissue contained within muscles while perimuscular adipose tissues (PAT) designates the remaining adipose tissue, mainly the fat deposits between the muscles.

In healthy subjects, fat is mainly present as SAT, whereas IAT is almost absent (**Figure 2A**). On the contrary, in neuromuscular diseases, muscle tissue is submitted to histological changes leading to a progressive replacement of muscle tissue by adipocytes. Intramuscular fat infiltration can even lead to muscle necrosis and fibrosis (**Figure 2D**).

Segmentation strategies have evolved over the years with improvements in image quality and clinical needs. The first approaches intended to separate muscle and fat deposits with no distinction between subcutaneous and internal compartment fat. Clinical research has revealed that perimuscular and intramuscular adipose tissue are not part of the same metabolic processes (21) but their respective contributions are not yet fully understood (22). Segmentation strategies have therefore evolved toward the segmentation of muscle regions or individual muscles in order to allow the precise quantification of each adipose compartments. This has also been facilitated by improved image quality, which has allowed better visualization of the boundaries between tissues. It should be noted that the nomenclature of the different adipose tissues is still not clearly defined in the research field of segmentation strategies. The terms "intermuscular fat" and "intramuscular fat" have been confused and misleadingly used in many of the articles we reviewed in this study.

2.3. Validation of Segmentation Approaches

Validation is a crucial step in the development of automatic segmentation methods. It intends to assess the effectiveness of an automatic segmentation method based on the comparative analysis between the automatic segmentation provided by an algorithm and a ground truth produced by one or more experts in the field, usually radiologists.

Several complementary metrics have been commonly used. They can assess the overlap between segmentations, the distance between segmentation contour points or the volumes computed from the segmentations. Each metric is actually sensitive to one type of segmentation error (size, location, and shape) and none can take all error types into account (23).

For the muscle segmentation methods reported in the present review, the most commonly used metrics are the relative volume difference (RVD), the dice similarity coefficient (DSC), the Hausdorff distance (HD), and the average surface distance (ASD).

Let X be the segmentation resulting from the algorithm and Y the ground truth. RVD is computed taking into account the volumes quantified from the manual and automatic segmentations. It actually refers to the ratio between the $|X - Y|$ difference and $|Y|$. The RVD score is a relevant metric in clinical fields related to muscle because many studies used muscle volume change as a biomarker. Nevertheless, this metric does not allow a geometric and spatial analysis between the manual and automatic segmentations.

The DSC measures the relative overlap between X and Y. It is calculated as the ratio of twice the intersection between X and Y by the number of combined elements of X and Y. As defined in (1), DSC can also be expressed in terms of true positive (TP), false negative (FN), and false positive (FP). The DSC values can range from 0 to 1, 1 indicating the largest similarity between segmentations.

$$\mathrm{DSC} = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2 \cdot \mathrm{TP}}{\mathrm{FN} + 2 \cdot \mathrm{TP} + \mathrm{FP}} \quad (1)$$

In addition to DSC, the distance between the segmentation boundaries can be computed in order to assess the segmentation robustness regardless of the volume. Let ∂X be the segmentation boundary of X and ∂Y be the boundary of Y. For HD and ASD, the smallest distance separating the boundaries is measured between each point of ∂X and ∂Y. The distance between two points x and y is the Euclidian distance $\delta(x, y) = ||x - y||$. The HD is calculated as the maximum of these distances (2) and the ASD as the average of the distances (3). Both distances are expressed in millimeters and low values are desirable for an accurate segmentation.

$$\mathrm{HD}(\partial X, \partial Y) = \max\{\sup_{y \in \partial Y} \inf_{x \in \partial X} \delta(x, y), \sup_{x \in \partial X} \inf_{y \in \partial Y} \delta(x, y)\} \quad (2)$$

$$\mathrm{ASD}(\partial X, \partial Y) = \frac{\sum_{x \in \partial X} \delta(x, \partial Y) + \sum_{y \in \partial Y} \delta(\partial X, y)}{|\partial X| + |\partial Y|} \quad (3)$$

The three geometric metrics, DSC, HD, and ASD, are complementary and should all be evaluated in order to properly

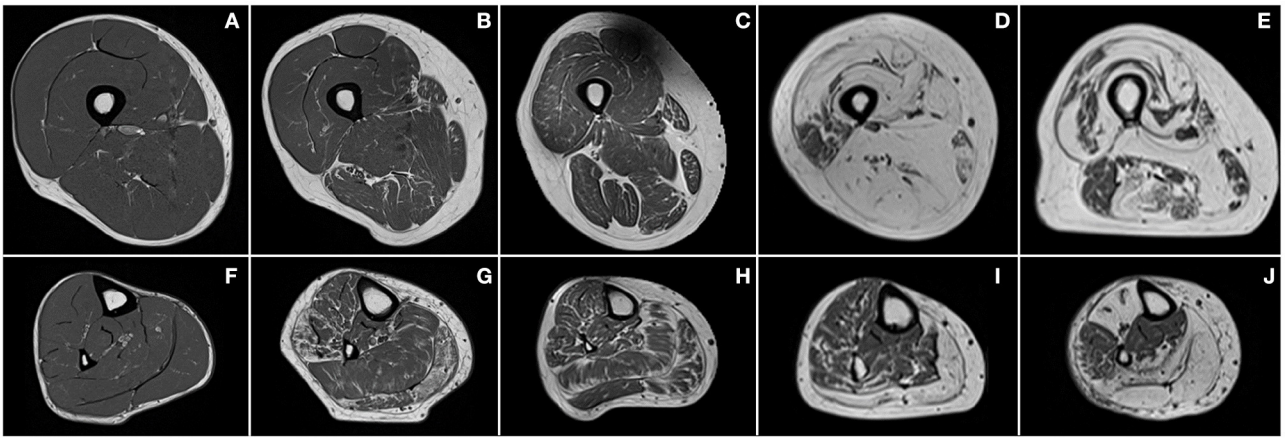


FIGURE 2 | Examples of T_1-weighted images of thighs (first row) and lower legs (second row) of a healthy subject **(A,F)** and patients with Charcot–Marie–Tooth disease type 1A **(B,G)**, myotonic dystrophy type 1 **(C,H)**, facioscapulohumeral muscular dystrophy **(D,I)**, and inclusion body myositis **(E,J)**.

validate a segmentation method. The DSC and ASD provide global information on the segmentation, i.e., the overlap and the distance between the boundaries, respectively. Unlike DSC and ASD, HD is very sensitive to outliers and to slight shape differences.

3. EVOLUTION OF SEGMENTATION STRATEGIES

3.1. Automatic Separation Between Muscle and Fat Deposits

Over the last 15 years, several automated methods have been reported in the literature with the aim of separating muscle and fat tissue in lower limbs MR images.

3.1.1. Separation Between Muscle and Adipose Tissue

Early automated segmentation methods took advantage of the contrast between fat, muscle, and bone tissues. Accordingly, Mattei et al. proposed a method for the semi-automatic segmentation of the muscle compartment of the thigh based on a histogram representation of T_1-weighted images (24). Based on user-defined thresholds for muscle and fat pixel intensities, the method was validated through a reproducibility study of the results between 3 experts but no metric related to the segmentation accuracy was reported. Becker et al. proposed an automatic method for the separation between muscle and fat using a series of thresholding, morphological, and connectivity operations enhanced by the use of the four different kind of images provided by the chemical-shift DIXON sequences (25). They reported a DSC score larger than 0.95 for muscle segmentation performed in seven slices selected along the lower limbs of four subjects.

Threshold-based methods are fast and simple to implement but are quite sensitive to noise, imaging artifacts, and above all need empirical thresholding. Automatic methods such as K-means clustering have been developed to address this issue (26). These methods classify tissues according to the intensity of each pixel. Other automatic methods based on Gaussian mixture model (GMM) histogram analysis have been reported (27). The corresponding results were better because of the unsupervised learning, which allows the algorithm to adapt to each image and makes it more robust.

Because K-means clustering approaches are based on the assumption that a feature vector belongs to only one class, they have been recognized as ill-suited for MR image segmentation when classes overlap or when the information is unclear and uncertain (28). Partial volume effects between muscle and fat near muscle boundaries and inter- or intra-muscular fat infiltration lead to class uncertainties. Fuzzy c-means (FCM) clustering algorithm has been developed in order to overcome this issue. An FCM clustering algorithm was proposed by Barra et al. to estimate 3D volumes of muscle and fat on thigh images (29). The method was reproducible with respect to volume estimates in five images acquired on the same day from three subjects but no comparison was performed with ground truth segmentations.

Methods based on clustering or histogram analysis allowed fat and muscle tissues to be distinguished but SAT and IAT remained undistinguishable. This is of importance considering that IAT is directly related to the pathological process of neuromuscular disorders, whereas SAT is not.

3.1.2. Separation Between Perimuscular and Subcutaneous Adipose Tissue

Since subcutaneous and internal adipose tissue have to be distinguished, several approaches have been proposed for the segmentation of the internal SAT border, in addition to fat and muscle tissue separations.

Although Valentinitsch et al. (30) applied K-means clustering on the different images resulting from chemical shift-based water-fat separation MR sequences successively, Yang et al. (31) proposed a machine learning algorithm using the whole set of images at a time. Both of these approaches allowed IAT to be distinguished from SAT using basic morphological operations such as dilatation, erosion, and connected components. These methods were assessed on a single chosen slice position for

Valentinish et al. and on a 3D stack of slices of thigh for Yang et al. For both methods, good results were obtained for the segmentation of muscular tissue in "healthy" images with DSC values higher than 0.94 and ASD around 0.80 mm. However, Valentinish et al. highlighted that this kind of approaches may not identify the correct muscle envelope if a muscle next to the SAT region is fatty infiltrated or surrounded by a substantial amount of fat, a common scenario in patients with neuromuscular disorders (**Figure 2**). Moreover, defining the delimitation between IAT and SAT as the muscular envelope was misleading given that the true natural boundary is the *fascia lata* (20), which may not appear close to the muscle.

To address the issue of segmenting the *fascia lata*, which is a very thin tissue, poorly contrasted, and partially invisible in MR images, a few methods based on snake, active geodesic contours (32) have been proposed.

This kind of algorithm may be able to perform an active contour evaluation toward weak edges, such as that of *fascia lata* (**Figure 1**). Snake algorithm was used in similar methods on T_1-weighted images by (33, 34) with a difference regarding contour initialization. Although Makrogiannis et al. used the leg boundary segmented by morphological operations, Orgiu et al. implemented a rough *fascia lata* segmentation defined by the muscle envelope segmented with an FCM clustering followed by morphological operations. Positano et al. used a gradient vector flow snakes (35), an extension of snake active contours, which does not need to be initialized close to the boundary and is able to converge to the boundary concavities (27). Succession of active contours initialized by a circle surrounding the leg were applied to segment the external SAT contour, the internal SAT contour, and finally the bone contour. External force of the snakes used edge map derived of a fat map previously created with an FCM clustering. Inside the internal border of the SAT, Positano et al. used a GMM approach to separate fat and muscle in the perimuscular region, whereas Makrogiannis et al. applied K-means clustering to the combined space of fat image and water image. These approaches were assessed in MR images of healthy volunteers (33) and obese patients (27, 34). No direct comparison was made with ground truth segmentation but good correlations for volume quantification between manual and automatic segmentations were reported for muscle and fat tissues within the internal SAT border. Orgiu et al. reported a mean ASD value of 0.81 mm for the *fascia lata* segmentation but no indication about the HD.

More recently, several approaches have been proposed to enhance active contour methods for *fascia lata* segmentation with learning methods and line detection filters.

A random forest approach coupled with sparsity constraints to fix the noise caused by veins was proposed by Tan et al. with the aim of learning a 2D *fascia lata* detector and incorporating it into the external energy terms of a gradient vector flow snake (36). They reported a high average DSC of 0.97 and an average ASD of 1.37 mm in thigh images of osteoarthritis patients. This method clearly outperformed those based on classic active contour model (37). In two others similar methods, geodesic active contours, also initialized with muscle envelope, were enhanced by line detection filter, which extracts *fascia lata* point candidates (38) and a vessel enhancement filtering, which distinguishes plate-like shapes (39). Kovacs et al. assayed their method on T_1-weighted images of myopathic patients and DSC for the segmentation of non-affected and mildly affected muscles was 0.93, whereas it was reduced (0.80) for severely affected muscles. These results can be explained by the volume dependency of the DSC and the fact that muscle volume was lower in severely affected patients. For the detection of the true muscle envelope (i.e., the *fascia lata*), HD scores were systematically high with an average of 13 mm regardless of the pathology severity. Chaudry et al. proposed a semi-automatic method based on live-wire to refine the *fascia lata* automatic delineation in addition of their automatic approach. Manual corrections were partially necessary for 40% of the datasets (23 healthy young men and 50 elderly sarcopenic men with a moderate level of fat deposits). No direct comparison with ground truth segmentations was performed.

Although the detection of the muscle envelope with intensity-based clustering or active contour approaches is ill-suited for images of patients with neuromuscular disorders, these methods have been used as an initialization step in most of the studies dedicated to *fascia lata* detection. As expected, segmentation of images with a severe fat infiltration was of poor quality. Other approaches therefore had to be proposed.

Chambers et al. (40) introduced a method based on a live-wire approach for muscle region segmentation. A fingerprint-based algorithm was used to overcome the limitations of basic live-wire approaches, which are sensitive to the additive noise and small textural information. The muscle region border was then identified using an exponential cost function related to the edge information. The internal border of the SAT was first detected by keeping the edges closest to the border of the leg. The procedure was performed on the slice with the largest cross-sectional area and then the border search was restricted to adjacent slices, assuming the location of the boundaries is fairly similar from one slice to another. The method was validated on 10 T_1-weighted images of facioscapulohumeral muscular dystrophy (FSHD) patients and the DSC and ASD values were 0.89 and 0.10 mm, respectively. They also demonstrated that the state-of-art FCM clustering and active contour methods were less robust than their method when fat infiltration was present.

Very recently, Gadermayr et al. (41) evaluated up-to-date approaches for the segmentation of the whole muscle region in datasets of patients with neuromuscular disorders. For mildly to moderately infiltrated patients, all the tested approaches allowed an accurate segmentation with mean DSC values above 0.90. For images with a severe fat infiltration, they demonstrated that a graph cut approach incorporating shape knowledge exhibited a more accurate segmentation than the other methods with an average DSC value of 0.80.

3.1.3. Toward Fat Infiltration Measurement

Overall, several studies have been proposed for the segmentation of muscle and internal adipose tissue in MR images of lower limbs. Only a few have been validated for images of patients with neuromuscular disorders with a rather limited number of images and metrics related to segmentation. One has to keep in mind that the initial clinical driving force related

to these segmentation methods was related to the automatic quantification of fat-unaffected muscle volume. Accordingly, Müller et al. used intensity-based segmentation approaches on T_1-weighted images in order to quantify the remaining muscular tissue. The corresponding metric allowed to distinguish myopathic and neuropathic patients from control subjects (42). Similar methods have been used for the estimation of fat infiltration in images of patients with neuromuscular disorders. This clinical purpose faced several issues. Lareau-Trudel et al. reported, in a clinical study, that separation between SAT and IAT with active contour methods failed for patients with severe fat infiltration and manual corrections had to be performed in 20% of the cohort (15). Furthermore, in many of the studies reviewed, authors claimed to propose a quantification of the intermuscular fat fraction while using ratios between fat-unaffected muscle volume and adipose tissue volume contained within the *fascia lata*. No separation between intermuscular and intramuscular fats was performed and could be hardly expected with such intensity-based approaches. The assessment of intramuscular adipose tissue is of utmost importance in neuromuscular diseases given that it is a hallmark of the disease process.

Overall, the segmentation approaches reviewed above are not adequate for the automatic quantification of fat infiltration, whereas automatic segmentation of individual muscles is warranted if one intends to thoroughly assess a pathological process.

3.2. Automatic Segmentation of Muscle Regions

The segmentation of individual muscles would have two main advantages. Fat infiltration and other pathological changes could be assessed in individual muscles. In addition, the processing of 3D datasets could be useful to investigate potential changes along the proximo-distal axis. Accordingly, fat infiltration patterns of individual muscles have been reported in a few clinical studies (5, 16).

Both the fat infiltration severity together with the time-dependent progression can largely vary between muscles in a given patient, or between patients with a given disease (**Figure 3**) and between various muscle disorders (**Figure 2**). On that basis, quantification of any MRI biomarker of interest in individual muscles and in a 3D dataset is of crucial importance so that cross-sectional as well as longitudinal studies and therapeutic trials could be properly performed. As most of the coming therapeutic strategies are more likely to halt or slow the disease progression rather than reversing the already established tissue damage, it may be worthwhile to exclude the observation of fully infiltrated muscles in order to pay a more particular attention to those muscles not fully infiltrated.

In addition, automated methods are necessary to study large cohorts of patients with neuromuscular diseases given that manual segmentation is not conceivable given the time required (9).

As mentioned above, the automated segmentation of individual muscles in MR images is challenging for multiple reasons. Muscles in an MR image of healthy subjects display similar intensities and textures so that they can be hardly distinguished. In addition, the boundaries between muscles are very thin and may be discontinuous or invisible in MR images. As illustrated in **Figure 2**, muscle boundaries can be even less visible in case of fat infiltration. Different muscle shapes and textures within and between patients also contribute to the challenge the automatic segmentation task. In addition, considering that muscles are made up of soft tissues, one can expect large shape changes due to external constraints imposed by leg and/or coil positioning in the MR scanner.

Over the last years, several studies have proposed automated approaches to overcome these difficulties and to provide accurate segmentations of muscles individually or grouped by regions. Most of the methods were based on shape-based approaches given that intensity-based approaches cannot distinguish individual muscles.

A wavelet-based encoding method was proposed by Essafi et al. (43) in order to provide a hierarchical encoding of shape variability. This approach was assessed for the segmentation of the *gastrocnemius medialis* in 20 healthy subjects and five patients. The corresponding results were of poor quality with an averaged DSC of 0.55.

Baudin et al. evaluated a method based on random walk in order to address the issue of incomplete contours, which may occur between muscles. Such a method relies on seeds positioning for the initiation step that can be done manually or automatically using atlases. This method has been evaluated in out-of-phase images, which have the particularity of showing strong contours between tissues and thus between individual muscles in images of healthy subjects. The corresponding DSC values were high, i.e., 0.80 ± 0.19 in 15 control subjects (44). The incorporation of shape prior knowledge and confidence map of muscle contours led to larger DSC values [0.86 ± 0.07; (45)]. Although such a method has not been assayed in patients with neuromuscular disorders, one could expect much lower DSC values given that the fat infiltration should erase the muscle borders. Andrews et al. (46) addressed this issue by designing a random forest boundary detector that seek to learn common appearances of the interfaces between muscles in order to distinguish them from intramuscular fat. This intermuscular boundary detector was combined with a statistical shape model over the space of generalized log-ratio representations and a pre-alignment approach based on GMM segmentation of the muscular tissue. They evaluated their method for the segmentation of individual thigh muscles in 10 healthy subjects and 10 patients with chronic obstructive pulmonary disease and they reported DSC values ranging from 0.70 ± 0.16 to 0.93 ± 0.06.

A method based on active contour model with an initialization through an active shape model was assessed for the segmentation of muscle regions in thigh of patients with knee osteoarthritis (47). The active shape model was trained in 113 axial MR slices for an assessment on 20 images. They reported that 50 training datasets were enough to obtain accurate segmentation. Good DSC values were reported for the segmentation of each muscle region. However, their method was only designed for the segmentation of a unique mid-slice. In addition, a manual

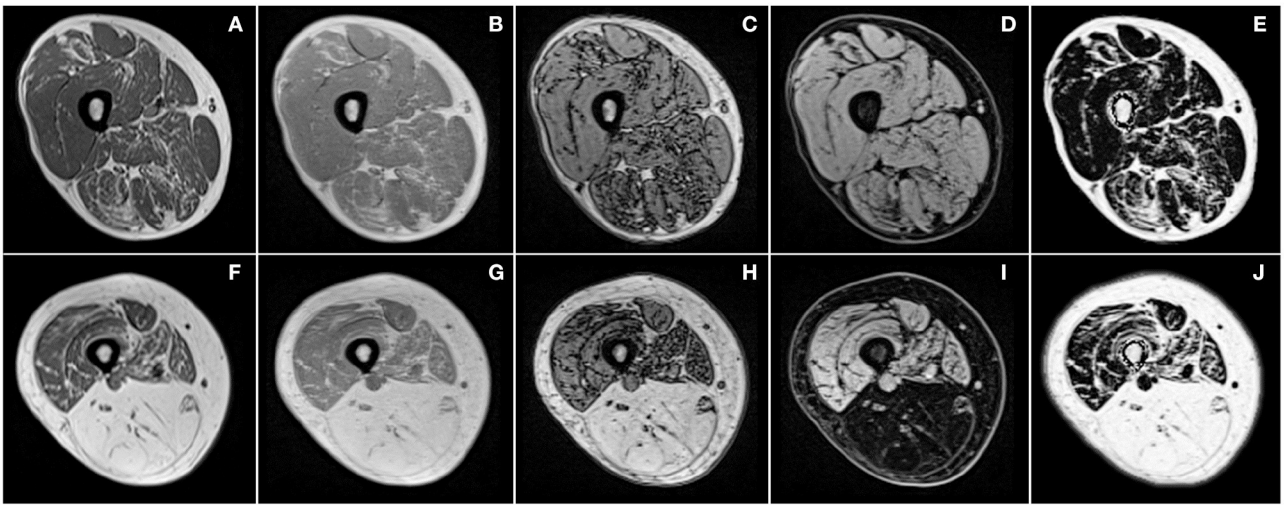


FIGURE 3 | Examples of T_1-weighted images **(A,F)** and images produced by chemical shift-based water-fat separation MR sequences for the thigh of two patients (first and second row) with different severity of facioscapulohumeral muscular dystrophy. Illustrations correspond to in-phase **(B,G)**, out-of-phase **(C,H)**, water **(D,I)** images, and quantitative fat fraction map **(E,J)**.

interaction of 3–5 min per slice was required for refining the initialization steps of both the active shape model and the active contour model.

3.2.1. Atlas-Based Approaches

Atlas-based approaches have been proposed for the automatic segmentation of individual muscles as they can incorporate spatial prior anatomical knowledge at individual muscles level. Atlas-based segmentation is a well-established concept (48) that has been widely applied to brain structures. These approaches treat segmentation as an image registration problem that aims at computing the optimal transformation fields from the pre-labeled atlases to the new image to be segmented. The atlas labels, once transferred to the new image domain, are merged and result in the final unique segmentation.

Such an approach has been used for the quantitative assessment of regional muscle volume in whole-body MR images (49). A multi-scaled and phase-based morphon method was selected for the registration because it would be less sensitive to MRI inhomogeneities. Water images were used since they display the least anatomical variation and contain the largest amount of information regarding muscle shapes. The multi-atlas process was performed twice, once with all the images of the atlas to obtain a first coarse segmentation, and then a second time using only the atlases having similar volumes to the coarse segmentations. Labels from atlases were merged into a probabilistic map and a threshold was set empirically to produce the final segmentation. Then, muscle and fat volumes were separated inside the segmented areas with a threshold. Using a leave-one-out strategy in 20 healthy subjects, good correlation was reported for muscle volume quantification. The mean TPVF was 93% for lower leg, posterior and anterior thigh compartments, and mean FPVF was 5% for lower leg and 8% for posterior and anterior thigh compartments. No geometric metrics related to the muscle envelope segmentation by multi-atlas process has been reported. In a recent study, this multi-atlas approach was also assessed as sensitive enough to detect significant changes in muscle volume following training activities (50).

Multi-atlas approaches are strongly dependent on the registration model used. Le Troter et al. (51) evaluated different registration methods from well-known open-source libraries to segment the four muscles of the *quadriceps femoris*. They demonstrated that multi-atlas process could be improved with initial registrations guided by the segmentation of SAT, muscle envelope, and bone, using a method described by Positano et al. (27). Similarly to Karlsson et al., the results for the segmentation of the whole *quadriceps femoris* in 25 healthy subjects were good with an averaged DSC score of 0.94 ± 0.03. The results were lower for the *vastus lateralis* and *rectus femoris* muscles, i.e., mean DSC values of 0.88 ± 0.08 and 0.84 ± 0.12. In addition, a 20% volume error with RVD scores of 0.17 ± 0.18 and 0.21 ± 0.24 were reported.

A multi-atlas approach based on B-spline nonrigid registrations has been reported by Belzunce et al. (52) with the aim of segmenting the gluteal muscles on hip and thigh images. Registrations were performed on muscle envelope pre-segmented with the Otsu algorithm. The approach was evaluated on both multi-atlas of T_1-weighted and DIXON in-phase images of 15 healthy subjects with DSC values of 0.94 ± 0.01 and RVD values of $1.5 \pm 4.3\%$. The assessment was only performed on the medial slices of the dataset excluding the extremities deemed to be more difficult to segment because of more variability and uncertainty.

Very recently, Mesbah et al. (53) introduced a Markov random field model combining appearance and spatial models with the prior shape information from atlases and so in order to segment

the three main muscle groups of the thigh. They reported good DSC scores (0.89 ± 0.05 to 0.95 ± 0.03) but the HD scores were of poor quality with an average ranging from 10.51 ± 6.37 to 31.53 ± 14.24 mm for the medial compartment. Furthermore, their method was assayed on images of healthy subjects and patients with spinal cord injury, for whom no fat infiltration occurs. Given that part of their method relied on intensity-based approach, it may be ill-suited for images of neuromuscular disorder patients. However, they demonstrated that their approach may outperformed those based on majority vote or STAPLE fusion following nonlinear atlas-based registrations.

The main advantage of multi-atlas methods is that any tissue compartment can be segmented according to a given atlas. This is of great interest for the study of neuromuscular diseases in which individual muscles are affected differently depending on the pathology. However, multi-atlas approaches have proposed mixed results for the segmentation of individual muscles and have only been validated in images of healthy subjects. In all the methods presented above, the final step of label fusion used only parts of the atlases, i.e., those closest to the image to be segmented. Various merging strategies have been used, i.e., majority vote (52) or more thoughtful algorithms like STEPS (51) or a Markov random field model (53). It should be kept in mind that parameters chosen empirically are optimized for a given training database and could have to be retuned for other images. On that basis, it may be necessary to have access to different subgroups of atlases that may be adapted to different types of images. This feature could be problematic for the generalization of these methods for the segmentation of pathological images considering the large between-muscles and between-subjects phenotypic variability. As neuromuscular diseases are classified as rare diseases, there is currently no database large enough to cover all the variability of pathologies as it may be available for the brain (54).

The above-reviewed approaches were designed for cross-sectional studies and a follow-up version has been originally reported by Le Troter et al. In a so-called single-atlas approach, they used the manual segmentation of a first time point as an atlas for the following time points (51). Since successive MR images of the same subject may show little anatomical variations, single atlas-based nonlinear registration can correctly transfer segmentation of the first time point images to the others. Assessed in healthy subject images, they reported much better results than the multi-atlas based approaches. DSC scores were above 0.89 for each individual muscle of the *quadriceps femoris* with RVD scores below 5%. A limitation of this method is that the full automatic aspect is lost with the need of an initial manually segmentation.

3.2.2. Toward Intramuscular Fat Infiltration Measurement

Although several automatic approaches have reported promising results for the segmentation of muscle regions in healthy subjects, none of them have been assessed for images of patients with neuromuscular disorders. To the best of our knowledge, only a single study has been devoted to the segmentation of muscles in patients with chronic obstructive pulmonary disease for whom fat infiltration was moderate and the approach was not accurate enough to be considered for clinical applications.

Overall, reliable automatic segmentation methods are still warranted for individual muscles in the context of neuromuscular pathologies. This is supportive of the conclusion from the review by Pons et al. (55). Although promising, the results already obtained in images of healthy subjects do not guarantee similar results on images of patients for whom factors such as fat infiltration may be problematic. Another critical issue is the availability of pre-labeled data sets of muscle MR images. The validation of a method requires manual segmentations from experts to be considered as the ground truth segmentations. As we previously mentioned, manual segmentation is time consuming and validation of methods has often been performed using a single slice (47).

The lack of automated segmentation methods is likely the reason why only a limited number of clinical studies have assessed qMRI scores at the individual muscles level in 3D datasets. Most of the studies have been performed considering a few slices only (4, 10) or a limited number of individual muscles (6).

3.3. Semi-Automatic Segmentation of Muscle Regions

As no automatic method has been validated for accurate segmentation of individual muscles in neuromuscular disorders, semi-automatic methods have been proposed to reach an optimized balance between segmentation accuracy and user's dedication.

A semi-automatic method for segmenting the whole *quadriceps femoris* was proposed with the manual delineation of a line separating this muscle group from the rest of the muscular envelope automatically pre-segmented with an adaptive threshold on T_1-weighted images (56). The method was assessed in healthy and elderly subjects and managed to reach a time saving of 87% with a mean DSC of 0.98. As we previously indicated, this kind of intensity-based methods is ill-suited for images of patients with neuromuscular disorders.

Two similar methods have been proposed to generate 3D muscle segmentations from manual segmentation on a limited number of 2D slices. One of the advantages of these approaches is that any tissue compartment can be segmented according to the manual segmentation defined on the few axial slices used as initialization of the process.

Jolivet et al. (57) proposed a model based on parametric-specific object method. Only a fast rough contouring of the muscle using polygons was necessary on the initial axial slices. Polygons were then matched to the muscle shape using an automatic contour optimization based on local gradient weighted by intensity similarity and distance to the rough contouring. Once segmentations were well-tuned on the initial slices, parametric-specific object was constructed and deformed to match the manual muscle segmentation. Process was iterative with the successive injection of the axial interpolated segmentation mask, after automatic contour optimization, in the parametric-specific object. The manual segmentation of

only 5 axial slices took 21 min (against 80 min for manual segmentation of all slices) and was enough to obtain an accurate 3D segmentation of all individual muscles with RVD scores lower than 5%. This method has not been tested in patients and one can expect that the automatic contour optimization may suffer from fat infiltration as it is a gradient-based approach.

The propagation of an initial manual segmentation to the remaining slices through a combination of nonlinear registration approaches has been originally reported by Ogier et al. The method takes advantage of the shape information from the initial manual segmentations with no other previous information regarding muscles shapes. It is also based on the anatomical information from a given image to the next in order to take into account the corresponding changes along the proximo-distal axis. On that basis, the initial segmentation could be automatically propagated along this axis. The initial manual segmentation has to be repeated each time a muscle was appearing or disappearing. The method was initially proposed and validated for the segmentation of the four muscles of the *quadriceps femoris* group in T_1-weighted images of 11 healthy subjects (58). Validation has been then extended for the segmentation of all individual thigh muscles in healthy subjects (59). Mean DSC scores of 0.90 ± 0.03 was reported with a manual input for 30% of the slices only. The semi-automatic method was also assessed for the segmentation of thigh and lower leg individual muscles in 10 patients with myotonic dystrophy type 1 (60). Using Dixon images recorded in both thigh and lower leg, only 9 out of 50 slices were manually segmented. Using water images, a mean DSC value of 0.91 ± 0.04 was reported and the results were similar regardless of the type of Dixon image used. In addition, an excellent reliability was also quantified from the comparative analysis between fat fraction computed from the segmented images and from the manual segmentation. The method has been deemed sufficiently robust for clinical applications and assayed in cross-sectional studies, which evaluated the pattern of fat infiltration in muscles in two different neuromuscular pathologies (12, 16).

Nonlinear registrations were also used for semi-automated segmentation methods dedicated to longitudinal studies.

Single-atlas approach was first assayed for the 3D segmentation of anterior and posterior thigh compartments in a 2-years follow-up study of patients with facioscapulohumeral muscular dystrophy, some at severe stages of intramuscular fat infiltration (61). No metrics of segmentation accuracy were reported but the fat quantification estimated through the semi-automatic segmentation showed a good reproducibility and repeatability as well as a good correlation with clinical scores.

Based on the single-atlas approach, Ogier et al. (60) reported a follow-up study in 10 myotonic dystrophy type 1 patients assessed twice 10 months apart. The 3D supervised segmentation of the first time point using the original semi-automatic method previously proposed (58) was used as a template for the automatic propagation to the second time point. The combined methods allowed an accurate segmentation with a DSC of 0.87 ± 0.07. Similarly to what has been obtained for the semi-automatic segmentation of baseline images, an excellent reliability was also observed between the fat fraction quantified from the automatic and manual segmentations. The combined methods provided the first complete framework dedicated to individual muscles segmentation and follow-up in patients with a neuromuscular disorder. Both the segmentation and a quantitative metric (fat fraction) in individual muscles were accurate, while the number of slices manually segmented was substantially reduced. The follow-up segmentation was performed with no additional manual segmentations and this could be translated for multiple repeated time points.

4. DEEP LEARNING-BASED SEGMENTATION METHODS

In the previous sections, evidence has been provided indicating that completely automatic methods are not efficient for a robust segmentation of individual fat-infiltrated muscles. On the contrary, semi-automatic methods can be robust and useable for clinical applications but the manual initialization remains an issue in terms of user dedication. Given the recent promising results reported in the field of medical imaging, deep learning-based segmentation methods are appealing (62).

Deep learning methods are part of machine learning methods, which have proven their efficiency in the diagnosis of neuromuscular diseases. Machine learning algorithm such as random forest models have been able to overcome experts for complicated diagnostic tasks (63). Deep learning methods are generally based on artificial neural networks, which are supervised to learn the segmentation process from manually segmented images provided as training examples. Neural networks rely on pixel intensities and image characteristics in order to compute the final segmentation. For images segmentation tasks, network architectures are built on the basis of convolutional encoder-decoder (CED) network. This kind of networks combines paired encoder and decoder networks and have the advantage of producing a result with a resolution similar to the initial images. This architecture can be seen as the association of a contracting path to capture the context and a symmetric expanding path that allows the image reconstruction. Various CED have been used for images segmentation. Among them, U-Net is considered as the standard CED architecture for image classification tasks (64) because of its efficient way of reconstructing the segmentation using information from the contracting path.

Similar to the conventional methods discussed in the previous sections, deep learning based segmentation methods have been applied for the different segmentation strategies, i.e., from the separation of muscle and fat deposits to the segmentation of individual muscles.

The very first deep learning approach applied on lower limb MR images was used in order to detect the *fascia lata*. Two studies intended to address this issue using a 5-layer network combined with a dual active contour model (65) or a U-Net architecture (66). Yao et al. used T_1-weighted images while Amer et al. showed the interest of combining T_2-weighted and PD images for their study. Both of them provided high-quality results with DSC values larger than 0.97 ± 0.02. Distinction

between adipose and healthy muscle tissue was performed using the same networks and the corresponding DSC values were also high, i.e., 0.91 (66) and 0.94 ± 0.07 (65) for muscle detection. Recently, impressive DSC scores of 0.97 were obtained with an improved U-Net structure using residual connections and dense blocks (67). However, such a classification did not allow to distinguish perimuscular and intramuscular adipose tissue. Using a U-Net architecture, Gadermayr et al. intended to segment healthy and fat-infiltrated muscle all-together on T_1-weighted images, allowing the distinction of intramuscular from perimuscular adipose tissue (68). Given the complexity of this task, corresponding DSC values were smaller (around 0.88 ± 0.05), illustrating a poorer segmentation quality.

With a similar purpose of distinguishing intramuscular and perimuscular fat, studies have been conducted with the aim of segmenting individual muscles. The AlexNet network was used by Ghosh et al. (69) to produce a principal component analysis of the segmentation, leading to poor results with average DSC of 0.85 ± 0.09. Standard deviation score illustrated the high variability of the results, which is known as a major weakness of deep learning methods (70). Better results were obtained with a U-Net architecture on T_1-weighted images (71) with average DSC reaching 0.95 ± 0.03. More recently, Ding et al. (72) used U-Net on fat-water decomposition MR images to segment 4 muscle regions and obtained DSC scores around 0.89 ± 0.03 in both healthy and affected subjects. However, as the network was trained with a single slice position, the high muscle shape variability along the proximo-distal axis could not be taken into consideration. One way to take into account the variability of the proximo-distal shape is to consider the muscle as a volume, which can be done using a 3D segmentation neural network. In this field, Conze et al. (73) demonstrated the interest of 3D segmentation for the segmentation of individual shoulder muscles. A limitation to 3D CED is the memory necessary to train it. Ni et al. (74) used bounding boxes around organs to reduce resolution and prevent memory growth. This method was applied with 3D U-Net on complete lower limb images of athletes, obtaining mean DSC on 35 muscles of 0.89 ± 0.03. A very recent study proposed to use an edge-aware network based on U-Net and reached a DSC of 0.90 ± 0.09 and an ASD of 1.37 ± 0.92 on both healthy and affected subjects (75).

Solutions to Scarcity of Data

The methods presented on the detection of *fascia lata*, the classification of adipose tissue, or the segmentation of individual muscles all showed promising results. However, all of them faced the problem of data availability due to the scarcity of annotated images of patients with pathological changes. Indeed, public datasets of annotated limbs MR images are scarce, unlike for other organs such as the brain (76).

An interesting solution for this issue relies on either the use of unlabeled data or the creation of artificial training examples. Amer et al. (66) proposed a semi-supervised learning method, which uses both labeled and unlabeled datasets. In that case, each image does not have to be annotated before the network training phase and one can increase the database without a human intervention for the labeling process. Anwar et al. (67) proposed to use a CED on unlabeled data to create labels and thus enlarge their dataset. However, unlabeled data are not always available especially for the study of rare diseases. For this purpose, methods based on data generation using neural networks are emerging since the founding article on generative adversarial networks (GAN) (77). Recently, Yi et al. (78) made a review regarding the application of such methods in medical imaging. For lower limb muscle segmentation, solutions based on GAN were assessed with the aim of generating pathological images (68). Many issues related to the realistic nature of the generated images and their variability have still to be addressed.

One has to keep in mind that the neural network training phase tightly relies on a tuning phase of the network hyper-parameters, which has to be empirically performed thereby reducing the fully automatic aspect of the method. In other words, a given network has to be optimized for a given dataset. Deep learning solutions have to be optimized by experts and this commonly takes hours of implementation. In addition, the training phases can be time consuming and the final results could be equivalent to those obtained with more conventional methods (53).

Overall, although deep learning tools for the segmentation of fat-infiltrated muscles have a great potential, one should keep in mind two major issues related to the availability of large amount of annotated data and the need of a specific network tuning for each dataset. Semi-automatic propagation methods have proven their efficiency (60) and could be used to annotate large amount of images. These methods could be combined to data augmentation for the generation of databases compatible with deep learning methods. The corresponding potential is still to be proven for the segmentation of fat-infiltrated muscles.

5. CONCLUSION

This review highlighted the lack of fully automated approaches that could produce accurate segmentations of muscle images of patients with neuromuscular disorders. The few validated methods that addressed the difficulty of segmenting images with severe infiltrated muscles were proposed for the whole muscle only. That might not be optimal in neuromuscular disorders in which individual muscles are seen to be affected differently. However, for segmentation of individual muscles, approaches that validated segmentations that were accurate enough for clinical use were evaluated only on healthy subjects. Specific studies are warranted for the extrapolation of these approaches to images of pathological muscles as the confounding factors differ. Indeed, the issue of distinguishing muscle, intramuscular adipose tissue, and subcutaneous adipose tissue seems to be crucial for the follow up of patients with a severe fat infiltration. Semi-automatic methods has proven some efficiency in clinical context. But even if they reduce the manual load required for the study of large cohorts, some manual interventions are still needed. As recent techniques, deep learning based approaches are promising but they need databases that are representative enough of typical neuromuscular disease images. The community should promote the emergence of common dedicated image databases.

AUTHOR CONTRIBUTIONS

AO, M-EB, and DB: guarantors of integrity of entire study. All authors: study concepts, manuscript drafting or manuscript revision for important intellectual content, approval of final version of submitted manuscript, agrees to ensure any questions related to the work are appropriately resolved, literature research, and manuscript editing.

REFERENCES

1. Mercuri E, Pichiecchio A, Allsop J, Messina S, Pane M, Muntoni F. Muscle MRI in inherited neuromuscular disorders: past, present, and future. *J Magn Reson Imaging*. (2007) 25:433–40. doi: 10.1002/jmri.20804
2. Janssen BH, Voet NB, Nabuurs CI, Kan HE, de Rooy JW, Geurts AC, et al. Distinct disease phases in muscles of facioscapulohumeral dystrophy patients identified by MR detected fat infiltration. *PLoS ONE*. (2014) 9:e85416. doi: 10.1371/journal.pone.0085416
3. Wokke B, Van Den Bergen J, Versluis M, Niks E, Milles J, Webb A, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. *Neuromusc Disord*. (2014) 24:409–16. doi: 10.1016/j.nmd.2014.01.015
4. Morrow JM, Sinclair CD, Fischmann A, Machado PM, Reilly MM, Yousry TA, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. *Lancet Neurol*. (2016) 15:65–77. doi: 10.1016/S1474-4422(15)00242-2
5. Heskamp L, van Nimwegen M, Ploegmakers MJ, Bassez G, Deux JF, Cumming SA, et al. Lower extremity muscle pathology in myotonic dystrophy type 1 assessed by quantitative MRI. *Neurology*. (2019) 92:e2803–14. doi: 10.1212/WNL.0000000000007648
6. Naarding KJ, Reyngoudt H, van Zwet EW, Hooijmans MT, Tian C, Rybalsky I, et al. MRI vastus lateralis fat fraction predicts loss of ambulation in Duchenne muscular dystrophy. *Neurology*. (2020) 94:e1386–94. doi: 10.1212/WNL.0000000000008939
7. Janssen B, Voet N, Geurts A, van Engelen B, Heerschap A. Quantitative MRI reveals decelerated fatty infiltration in muscles of active FSHD patients. *Neurology*. (2016) 86:1700–7. doi: 10.1212/WNL.0000000000002640
8. Heskamp L, Okkersen K, van Nimwegen M, Ploegmakers MJ, Bassez G, Deux JF, et al. Quantitative muscle MRI depicts increased muscle mass after a behavioral change in myotonic dystrophy type 1. *Radiology*. (2020) 2020:192518. doi: 10.1148/radiol.2020192518
9. Barnouin Y, Butler-Browne G, Voit T, Reversat D, Azzabou N, Leroux G, et al. Manual segmentation of individual muscles of the quadriceps femoris using MRI: a reappraisal. *J Magn Reson Imaging*. (2014) 40:239–47. doi: 10.1002/jmri.24370
10. Carlier PG, Marty B, Scheidegger O, Loureiro de Sousa P, Baudin PY, Snezhko E, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging and spectroscopy as an outcome measure for clinical trials. *J Neuromusc Dis*. (2016) 3:1–28. doi: 10.3233/JND-160145
11. Hooijmans M, Niks E, Burakiewicz J, Anastasopoulos C, van den Berg S, van Zwet E, et al. Non-uniform muscle fat replacement along the proximodistal axis in Duchenne muscular dystrophy. *Neuromusc Disord*. (2017) 27:458–64. doi: 10.1016/j.nmd.2017.02.009
12. Ansari B, Salort-Campana E, Ogier A, Le Troter PhD A, De Sainte Marie B, Guye M, et al. Quantitative muscle MRI study of patients with sporadic inclusion body myositis. *Muscle Nerve*. (2020) 61:496–503. doi: 10.1002/mus.26813
13. Chrzanowski SM, Baligand C, Willcoks RJ, Deol J, Schmalfuss I, Lott DJ, et al. Multi-slice MRI reveals heterogeneity in disease distribution along the length of muscle in Duchenne muscular dystrophy. *Acta Myol*. (2017) 36:151.
14. Kornblum C, Lutterbey G, Bogdanow M, Kesper K, Schild H, Schröder R, et al. Distinct neuromuscular phenotypes in myotonic dystrophy types 1 and 2. *J Neurol*. (2006) 253:753–61. doi: 10.1007/s00415-006-0111-5
15. Lareau-Trudel E, Le Troter A, Ghattas B, Pouget J, Attarian S, Bendahan D, et al. Muscle quantitative MR imaging and clustering analysis in patients with facioscapulohumeral muscular dystrophy type 1. *PLoS ONE*. (2015) 10:e0132717. doi: 10.1371/journal.pone.0132717
16. Bas J, Ogier AC, Le Troter A, Delmont E, Leporq B, Pini L, et al. Fat fraction distribution in lower limb muscles of patients with CMT1A: a quantitative MRI study. *Neurology*. (2020) 94:e1480–7. doi: 10.1212/WNL.0000000000009013
17. Brunner G, Nambi V, Yang E, Kumar A, Virani SS, Kougias P, et al. Automatic quantification of muscle volumes in magnetic resonance imaging scans of the lower extremities. *Magn Reson Imaging*. (2011) 29:1065–75. doi: 10.1016/j.mri.2011.02.033
18. Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. (2010) 29:1310–20. doi: 10.1109/TMI.2010.2046908
19. Dixon WT. Simple proton spectroscopic imaging. *Radiology*. (1984) 153:189–94 doi: 10.1148/radiology.153.1.6089263
20. Shen W, Wang Z, Punyanita M, Lei J, Sinav A, Kral JG, et al. Adipose tissue quantification by imaging methods: a proposed classification. *Obes Res*. (2003) 11:5–16. doi: 10.1038/oby.2003.3
21. Addison O, Marcus RL, LaStayo PC, Ryan AS. Intermuscular fat: a review of the consequences and causes. *Int J Endocrinol*. (2014) 2014:309570. doi: 10.1155/2014/309570
22. Ogawa M, Lester R, Akima H, Gorgey AS. Quantification of intermuscular and intramuscular adipose tissue using magnetic resonance imaging after neurodegenerative disorders. *Neural Regener Res*. (2017) 12:2100. doi: 10.4103/1673-5374.221170
23. Yeghiazaryan V, Voiculescu ID. Family of boundary overlap metrics for the evaluation of medical image segmentation. *J Med Imaging*. (2018) 5:015006. doi: 10.1117/1.JMI.5.1.015006
24. Mattei J, Fur YL, Cuge N, Guis S, Cozzone P, Bendahan D. Segmentation of fascias, fat and muscle from magnetic resonance images in humans: the DISPIMAG software. *Magn Reson Mater Phys Biol Med*. (2006) 19:275–9. doi: 10.1007/s10334-006-0051-1
25. Becker M, Magnenat-Thalmann N. Muscle tissue labeling of human lower limb in multi-channel mDixon MR imaging: concepts and applications. *IEEE/ACM Trans Comput Biol Bioinformatics*. (2015) 14:290–9. doi: 10.1109/TCBB.2015.2459679
26. Gray C, MacGillivray TJ, Eeley C, Stephens NA, Beggs I, Fearon KC, et al. Magnetic resonance imaging with k-means clustering objectively measures whole muscle volume compartments in sarcopenia/cancer cachexia. *Clin Nutr*. (2011) 30:106–11. doi: 10.1016/j.clnu.2010.07.012
27. Positano V, Christiansen T, Santarelli MF, Ringgaard S, Landini L, Gastaldelli A. Accurate segmentation of subcutaneous and intermuscular adipose tissue from MR images of the thigh. *J Magn Reson Imaging*. (2009) 29:677–84. doi: 10.1002/jmri.21699
28. Di Gesú V, Romeo L. An application of integrated clustering to MRI segmentation. *Pattern Recogn Lett*. (1994) 15:731–8. doi: 10.1016/0167-8655(94)90078-7
29. Barra V, Boire JY. Segmentation of fat and muscle from MR images of the thigh by a possibilistic clustering algorithm. *Comput Methods Prog Biomed*. (2002) 68:185–93. doi: 10.1016/S0169-2607(01)00172-9
30. Valentinitsch A, Karampinos D, Alizai H, Subburaj K, Kumar D, Link T, et al. Automated unsupervised multi-parametric classification of adipose tissue depots in skeletal muscle. *J Magn Reson Imaging*. (2013) 37:917–27. doi: 10.1002/jmri.23884
31. Yang YX, Chong MS, Tay L, Yew S, Yeo A, Tan CH. Automated assessment of thigh composition using machine learning for Dixon magnetic resonance images. *Magn Reson Mater Phys Biol Med*. (2016) 29:723–31. doi: 10.1007/s10334-016-0547-2
32. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. *Int J Comput. Vis*. (1988) 1:321–31. doi: 10.1007/BF00133570
33. Makrogiannis S, Serai S, Fishbein KW, Schreiber C, Ferrucci L, Spencer RG. Automated quantification of muscle and fat in the thigh from water-, fat-, and nonsuppressed MR images. *J Magn Reson Imaging*. (2012) 35:1152–61. doi: 10.1002/jmri.22842

34. Orgiu S, Lafortuna CL, Rastelli F, Cadioli M, Falini A, Rizzo G. Automatic muscle and fat segmentation in the thigh from T1-Weighted MRI. *J Magn Reson Imaging*. (2016) 43:601–10. doi: 10.1002/jmri.25031
35. Xu C, Prince JL. Snakes, shapes, and gradient vector flow. *IEEE Trans Image Process*. (1998) 7:359–69. doi: 10.1109/83.661186
36. Tan C, Yan Z, Yang D, Li K, Yu HJ, Engelke K, et al. Accurate thigh inter-muscular adipose quantification using a data-driven and sparsity-constrained deformable model. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)* Brooklyn, NY. (2015). p. 1130–4. doi: 10.1109/ISBI.2015.7164071
37. Tan C, Yan Z, Zhang S, Belaroussi B, Yu HJ, Miller C, et al. An automated and robust framework for quantification of muscle and fat in the thigh. In: *2014 22nd International Conference on Pattern Recognition* Stockholm. (2014). p. 3173–8. doi: 10.1109/ICPR.2014.547
38. Kovacs W, Liu CY, Summers R, Yao J. Identification of muscle and subcutaneous and intermuscular adipose tissue on thigh MRI of muscular dystrophy. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)* Prague. (2016). p. 176–9. doi: 10.1109/ISBI.2016.7493238
39. Chaudry O, Friedberger A, Grimm A, Uder M, Nagel AM, Kemmler W, et al. Segmentation of the fascia lata and reproducible quantification of intermuscular adipose tissue (IMAT) of the thigh. *Magn Reson Mater Phys Biol Med*. (2020) 1–10. doi: 10.1007/s10334-020-00878-w. [Epub ahead of print].
40. Chambers O, Milenković J, Pražnikar A, Tasič JF. Computer-based assessment for facioscapulohumeral dystrophy diagnosis. *Comput Methods Prog Biomed*. (2015) 120:37–48. doi: 10.1016/j.cmpb.2015.03.006
41. Gadermayr M, Disch C, Müller M, Merhof D, Gess B. A comprehensive study on automated muscle segmentation for assessing fat infiltration in neuromuscular diseases. *Magn Reson Imaging*. (2018) 48:20–6. doi: 10.1016/j.mri.2017.12.014
42. Müller M, Dohrn MF, Romanzetti S, Gadermayr M, Reetz K, Krämer NA, et al. Semi-automated volumetry of MRI serves as a biomarker in neuromuscular patients. *Muscle Nerve*. (2020) 61:600–7. doi: 10.1002/mus.26827
43. Essafi S, Langs G, Deux JF, Rahmouni A, Bassez G, Paragios N. Wavelet-driven knowledge-based MRI calf muscle segmentation. In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* Boston, MA. (2009). p. 225–8. doi: 10.1109/ISBI.2009.5193024
44. Baudin PY, Azzabou N, Carlier PG, Paragios N. Automatic skeletal muscle segmentation through random walks and graph-based seed placement. In: *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* Barcelona. (2012). p. 1036–9. doi: 10.1109/ISBI.2012.6235735
45. Baudin PY, Azzabou N, Carlier PG, Paragios N. Prior knowledge, random walks and human skeletal muscle segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Nice: Springer (2012). p. 569–76. doi: 10.1007/978-3-642-33415-3_70
46. Andrews S, Hamarneh G. The generalized log-ratio transformation: learning shape and adjacency priors for simultaneous thigh muscle segmentation. *IEEE Trans Med Imaging*. (2015) 34:1773–87. doi: 10.1109/TMI.2015.2403299
47. Kemnitz J, Eckstein F, Culvenor AG, Ruhdorfer A, Dannhauer T, Ring-Dimitriou S, et al. Validation of an active shape model-based semi-automated segmentation algorithm for the analysis of thigh muscle and adipose tissue cross-sectional areas. *Magn Reson Mater Phys Biol Med*. (2017) 30:489–503. doi: 10.1007/s10334-017-0622-3
48. Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: a survey. *Med Image Anal*. (2015) 24:205–19. doi: 10.1016/j.media.2015.06.012
49. Karlsson A, Rosander J, Romu T, Tallberg J, Grönqvist A, Borga M, et al. Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. *J Magn Reson Imaging*. (2015) 41:1558–69. doi: 10.1002/jmri.24726
50. Mandic M, Rullman E, Widholm P, Lilja M, Leinhard D, Thomas G, et al. Automated assessment of regional muscle volume and hypertrophy using MRI. *Sci Rep*. (2020) 10:2239 doi: 10.1038/s41598-020-59267-x
51. Le Troter A, Fouré A, Guye M, Confort-Gouny S, Mattei JP, Gondin J, et al. Volume measurements of individual muscles in human quadriceps femoris using atlas-based segmentation approaches. *Magn Reson Mater Phys Biol Med*. (2016) 29:245–57. doi: 10.1007/s10334-016-0535-6
52. Belzunce MA, Henckel J, Fotiadou A, Di Laura A, Hart A. Automated multi-atlas segmentation of gluteus maximus from Dixon and T1-weighted magnetic resonance images. *Magn Reson Mater Phys Biol Med*. (2020) 33:677–88. doi: 10.1007/s10334-020-00839-3
53. Mesbah S, Shalaby AM, Stills S, Soliman AM, Willhite A, Harkema SJ, et al. Novel stochastic framework for automatic segmentation of human thigh MRI volumes and its applications in spinal cord injured individuals. *PLoS ONE*. (2019) 14:e0216487. doi: 10.1371/journal.pone.0216487
54. Marcus DS, Wang TH, Parker J, Csernansky JG, Morris JC, Buckner RL. Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J Cogn Neurosci*. (2007) 19:1498–507. doi: 10.1162/jocn.2007.19.9.1498
55. Pons C, Borotikar B, Garetier M, Burdin V, Ben Salem D, Lempereur M, et al. Quantifying skeletal muscle volume and shape in humans using MRI: a systematic review of validity and reliability. *PLoS ONE*. (2018) 13:e0207847. doi: 10.1371/journal.pone.0207847
56. Ahmad E, Yap MH, Degens H, McPhee JS. Atlas-registration based image segmentation of MRI human thigh muscles in 3D space. In: *Medical Imaging 2014: Image Perception, Observer Performance, and Technology Assessment*. Vol. 9037. San Diego, CA: International Society for Optics and Photonics (2014). p. 90371L. doi: 10.1117/12.2043606
57. Jolivet E, Dion E, Rouch P, Dubois G, Charrier R, Payan C, et al. Skeletal muscle segmentation from MRI dataset using a model-based approach. *Comput Methods Biomech Biomed Eng*. (2014) 2:138–45. doi: 10.1080/21681163.2013.855146
58. Ogier A, Sdika M, Foure A, Le Troter A, Bendahan D. Individual muscle segmentation in MR images: a 3D propagation through 2D non-linear registration approaches. In: *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* Jeju Island. (2017). p. 317–20. doi: 10.1109/EMBC.2017.8036826
59. Secondulfo L, Ogier AC, Monte JR, Aengevaeren VL, Bendahan D, Nederveen AJ, et al. Supervised segmentation framework for evaluation of diffusion tensor imaging indices in skeletal muscle. *NMR in Biomed*. (2020) 34:e4406. doi: 10.1002/nbm.4406
60. Ogier AC, Heskamp L, Michel CP, Fouré A, Bellemare ME, Le Troter A, et al. A novel segmentation framework dedicated to the follow-up of fat infiltration in individual muscles of patients with neuromuscular disorders. *Magn Reson Med*. (2020) 83:1825–36. doi: 10.1002/mrm.28030
61. Fatehi F, Salort-Campana E, Le Troter A, Lareau-Trudel E, Bydder M, Fouré A, et al. Long-term follow-up of MRI changes in thigh muscles of patients with Facioscapulohumeral dystrophy: a quantitative study. *PLoS ONE*. (2017) 12:e0183825. doi: 10.1371/journal.pone.0183825
62. Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Zeitsch Med Phys*. (2019) 29:102–27. doi: 10.1016/j.zemedi.2018.11.002
63. Verdú-Díaz J, Alonso-Pérez J, Nu nez-Peralta C, Tasca G, Vissing J, Straub V, et al. Accuracy of a machine learning muscle MRI-based tool for the diagnosis of muscular dystrophies. *Neurology*. (2020) 94:e1094–102. doi: 10.1212/WNL.0000000000009068
64. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich: Springer (2015). p. 234–41. doi: 10.1007/978-3-319-24574-4_28
65. Yao J, Kovacs W, Hsieh N, Liu CY, Summers RM. Holistic segmentation of intermuscular adipose tissues on thigh MRI. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Quebec City, QC: Springer (2017). p. 737–45. doi: 10.1007/978-3-319-66182-7_84
66. Amer R, Nassar J, Bendahan D, Greenspan H, Ben-Eliezer N. Automatic segmentation of muscle tissue and inter-muscular fat in thigh and calf MRI images. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Shenzhen: Springer (2019). p. 219–27. doi: 10.1007/978-3-030-32245-8_25
67. Anwar SM, Irmakci I, Torigian DA, Jambawalikar S, Papadakis GZ, Akgun C, et al. Semi-supervised deep learning for multi-tissue segmentation from multi-contrast MRI. *J Signal Process Syst*. (2020) 1–14. doi: 10.1007/s11265-020-01612-4
68. Gadermayr M, Li K, Müller M, Truhn D, Krämer N, Merhof D, et al. Domain-specific data augmentation for segmenting MR images of fatty

infiltrated human thighs with neural networks. *J Magn Reson Imaging*. (2019) 49:1676–83. doi: 10.1002/jmri.26544
69. Ghosh S, Ray N, Boulanger P. A structured deep-learning based approach for the automated segmentation of human leg muscle from 3d MRI. In: *2017 14th Conference on Computer and Robot Vision (CRV)* Ottawa, ON. (2017). p. 117–23. doi: 10.1109/CRV.2017.32
70. Renard F, Guedria S, De Palma N, Vuillerme N. Variability and reproducibility in deep learning for medical image segmentation. *Sci Rep*. (2020) 10:1–16. doi: 10.1038/s41598-020-69920-0
71. Kemnitz J, Baumgartner CF, Eckstein F, Chaudhari A, Ruhdorfer A, Wirth W, et al. Clinical evaluation of fully automated thigh muscle and adipose tissue segmentation using a U-Net deep learning architecture in context of osteoarthritic knee pain. *Magn Reson Mater Phys Biol Med*. (2020) 33:483–93. doi: 10.1007/s10334-019-00816-5
72. Ding J, Cao P, Chang HC, Gao Y, Chan SHS, Vardhanabhuti V. Deep learning-based thigh muscle segmentation for reproducible fat fraction quantification using fat-water decomposition MRI. *Insights Imaging*. (2020) 11:1–11. doi: 10.1186/s13244-020-00946-8
73. Conze PH, Brochard S, Burdin V, Sheehan FT, Pons C. Healthy versus pathological learning transferability in shoulder muscle MRI segmentation using deep convolutional encoder-decoders. *Comput Med Imaging Graph*. (2020) 2020:101733. doi: 10.1016/j.compmedimag.2020.101733
74. Ni R, Meyer CH, Blemker SS, Hart JM, Feng X. Automatic segmentation of all lower limb muscles from high-resolution magnetic resonance imaging using a cascaded three-dimensional deep convolutional neural network. *J Med Imaging*. (2019) 6:044009. doi: 10.1117/1.JMI.6.4.044009
75. Guo Z, Zhang H, Chen Z, van der Plas E, Gutmann L, Thedens D, et al. Fully automated 3D segmentation of MR-imaged calf muscle compartments: neighborhood relationship enhanced fully convolutional network. *Comput Med Imaging Graph*. (2020) 2020:101835. doi: 10.1016/j.compmedimag.2020.101835
76. Akkus Z, Galimzianova A, Hoogi A, Rubin DL, Erickson BJ. Deep learning for brain MRI segmentation: state of the art and future directions. *J Digit Imaging*. (2017) 30:449–59. doi: 10.1007/s10278-017-9983-4
77. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. *Adv Neural Inform Process Syst*. (2014) 27:2672–80. doi: 10.5555/2969033.2969125
78. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med Image Anal*. (2019) 58:101552. doi: 10.1016/j.media.2019.101552

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 25 March 2021
doi: 10.3389/fneur.2021.637198

Normalization of Spinal Cord Total Cross-Sectional and Gray Matter Areas as Quantified With Radially Sampled Averaged Magnetization Inversion Recovery Acquisitions

Eva M. Kesenheimer[1,2,3], *Maria Janina Wendebourg*[1,2,3], *Matthias Weigel*[1,3,4], *Claudia Weidensteiner*[4,5], *Tanja Haas*[4], *Laura Richter*[1,2,3], *Laura Sander*[1,2,3], *Antal Horvath*[5], *Muhamed Barakovic*[3,5], *Philippe Cattin*[5], *Cristina Granziera*[1,3,5], *Oliver Bieri*[4,5] *and Regina Schlaeger*[1,2,3]*

[1] Neurologic Clinic and Policlinic, University Hospital Basel, Basel, Switzerland, [2] Department of Clinical Research, University of Basel, Basel, Switzerland, [3] Translational Imaging in Neurology (ThINK) Basel, Department of Medicine and Biomedical Engineering, Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel and University of Basel, Basel, Switzerland, [4] Division of Radiological Physics, Department of Radiology, University Hospital Basel, Basel, Switzerland, [5] Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland

OPEN ACCESS

Edited by:
Anna Pichiecchio,
Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy

Reviewed by:
Giovanni Savini,
University of Milan, Italy
Valentina Panara,
Università degli Studi G. d'Annunzio Chieti e Pescara, Italy

***Correspondence:**
Regina Schlaeger
regina.schlaeger@usb.ch

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: 02 December 2020
Accepted: 05 February 2021
Published: 25 March 2021

Citation:
Kesenheimer EM, Wendebourg MJ, Weigel M, Weidensteiner C, Haas T, Richter L, Sander L, Horvath A, Barakovic M, Cattin P, Granziera C, Bieri O and Schlaeger R (2021) Normalization of Spinal Cord Total Cross-Sectional and Gray Matter Areas as Quantified With Radially Sampled Averaged Magnetization Inversion Recovery Acquisitions. Front. Neurol. 12:637198. doi: 10.3389/fneur.2021.637198

Background: MR imaging of the spinal cord (SC) gray matter (GM) at the cervical and lumbar enlargements' level may be particularly informative in lower motor neuron disorders, e. g., spinal muscular atrophy, but also in other neurodegenerative or autoimmune diseases affecting the SC. Radially sampled averaged magnetization inversion recovery acquisition (rAMIRA) is a novel approach to perform SC imaging in clinical settings with favorable contrast and is well-suited for SC GM quantitation. However, before applying rAMIRA in clinical studies, it is important to understand (i) the sources of inter-subject variability of total SC cross-sectional areas (TCA) and GM area (GMA) measurements in healthy subjects and (ii) their relation to age and sex to facilitate the detection of pathology-associated changes. In this study, we aimed to develop normalization strategies for rAMIRA-derived SC metrics using skull and spine-based metrics to reduce anatomical variability.

Methods: Sixty-one healthy subjects (age range 11–93 years, 37.7% women) were investigated with axial two-dimensional rAMIRA imaging at 3T MRI. Cervical and thoracic levels including the level of the cervical (C4/C5) and lumbar enlargements (T_{max}) were examined. SC T2-weighted sagittal images and high-resolution 3D whole-brain T1-weighted images were acquired. TCA and GMAs were quantified. Anatomical variables with associations of $|r| > 0.30$ in univariate association with SC areas, and age and sex were used to construct normalization models using backward selection with $TCA_{C4/C5}$ as outcome. The effect of the normalization was assessed by % relative standard deviation (RSD) reductions.

Results: Mean inter-individual variability and the SD of the SC area metrics were considerable: $TCA_{C4/5}$: 8.1%/9.0; TCA_{Tmax}: 8.9%/6.5; $GMA_{C4/C5}$: 8.6%/2.2; GMA_{Tmax}: 12.2%/3.8. Normalization based on sex, brain WM volume, and spinal canal area resulted

in RSD reductions of 23.7% for TCAs and 12.0% for GM areas at C4/C5. Normalizations based on the area of spinal canal alone resulted in RSD reductions of 10.2% for TCAs and 9.6% for GM areas at C4/C5, respectively.

Discussion: Anatomic inter-individual variability of SC areas is substantial. This study identified effective normalization models for inter-subject variability reduction in TCA and SC GMA in healthy subjects based on rAMIRA imaging.

Keywords: spinal cord gray matter imaging, MRI, normalization, inter-subject variability, minors, spinal muscular atrophy

INTRODUCTION

Substantial advances in understanding spinal muscular atrophy (SMA) etiopathogenesis have catalyzed the development of novel therapeutic strategies. With the approval of the first disease-modifying treatments for SMA, the need for biomarkers that allow reliable monitoring of the disease course and the therapeutic response in SMA patients has substantially grown. Current advances in morphometric MRI development allow gray (GM) and white matter (WM) quantification in the spinal cord (SC) (1–7), which may help in improving the *in vivo* characterization of motor neuron diseases or other neurodegenerative SC diseases. Imaging the cervical and lumbar enlargements could be informative, especially in lower motor neuron diseases, e.g., SMA, or lower motor neuron-predominant amyotrophic lateral sclerosis (ALS).

Radially sampled averaged magnetization inversion recovery acquisition (rAMIRA) (8, 9) is a novel magnetic resonance imaging (MRI) approach to perform SC imaging with favorable contrast in clinical settings, which is well-suited for GM/WM quantitation not only in the cervical, but also in the thoracic SC. Briefly, after an inversion recovery preparation, rAMIRA typically acquires five radially sampled images with increasing inversion times (8). The first images of the series with shorter inversion times display high gray matter to white matter contrast, while the images with longer inversion times show a bright CSF in contrast to a dark SC (8, 9). These acquired inversion images can be combined to fine-tune and even enhance the signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) (9). Due to the radial sampling scheme with a balanced steady-state free precession readout module, rAMIRA provides a low sensitivity to motion effects such as heartbeat and breathing, which is a crucial issue in imaging of the thoracic SC (8). Based on these advantages and a good in-plane resolution, rAMIRA images are well-suited for quantifying both GM area and total cross-sectional area (TCA) in the SC.

More recently, several semi-automated and automated tools (10–17) have been developed for segmentation of the SC GM and WM from MR images, including promising automated segmentation algorithms specifically developed for the AMIRA approach (7, 14). The reliability of the segmentation methods in single center studies is in general high, and some methods have been validated in multi-centric settings (18). Despite these substantial advances, the anatomic accuracy for delineation of SC GM is still judged based on manual algorithms (18), in particular in studies involving the thoracic SC.

Prior to applying morphometric SC techniques such as rAMIRA in larger clinical studies in lower motor neuron disorders, it is important to understand the sources of inter-subject variability of SC GM and WM area measurements in healthy subjects to increase both the sensitivity and specificity in detecting pathology-associated changes. Normalization aims to reduce biological, anatomical variation unrelated to the disease.

Previous studies in multiple sclerosis patients proposed to decrease anatomic inter-subject variability by applying normalization approaches based on correlations between the upper cervical total cross-sectional SC area/cervical SC volume and (i) skull size in healthy subjects (19, 20), (ii) lumbar enlargement SC area (21), or (iii) SC length (22, 23). However, results remain partly conflicting. Papinutto et al. (24) recently reported a 17.7% reduction of inter-subject variability in upper cervical total cross-sectional SC area at the intervertebral disc level C2/C3 based on a normalization approach combining SIENAX v-scale (25) and the product of the maximum axial anterior–posterior and lateral diameters of the cervical spinal canal in a cohort of healthy middle-aged adults based on phase-sensitive inversion recovery imaging.

The level C2/C3 has been the major target in multiple sclerosis imaging studies (26, 27) as it is clearly situated above the cervical SC enlargement and is therefore anatomically less variable than the levels below.

Nevertheless, the lower levels of the cervical and thoracic SC that contain the motor neurons to the arm and leg muscles are of special relevance to the study of lower motor neuron disorders (e.g., SMA) and have been less well-studied. There are only a few MRI studies assessing the SC of children and adolescents (28–30), none focusing on the SC GM. Children and adolescents are the leading target group for the recently approved SMA treatments; therefore, data on SC GM variations are needed to develop treatment monitoring methods.

The aims of this study were to assess the anatomic inter-subject variability in TCA and GM areas at several levels in the cervical and thoracic SC based on rAMIRA imaging in a cohort of healthy subjects with a broad age range including both adults as well as minors and to develop and evaluate potential normalization strategies for inter-subject variability reduction.

METHODS

Participants

Sixty-one healthy subjects (range 11–93 years, mean age 46.0 years, SD 24.7, 37.7% women) including 18 minors (range 11–17 years, mean age 13.9, SD 1.9, 46.2% female) without a neurological or cognitive disease were included into the study after written informed consent was obtained. The local ethics committee approved the study.

MRI Acquisition

All participants were examined with the same 3T whole-body MR system (Siemens Magnetom PRISMA, Siemens Healthineers, Erlangen, Germany) using a 64-channel head and neck coil and the built-in spine coil for reception. Axial two-dimensional rAMIRA images (8, 9) were acquired perpendicular to the SC at the intervertebral disc levels C2/C3, C3/C4, C4/C5, C5/6, T9/T10, and T_{max} [level of the lumbar enlargement, which was identified by visual inspection on the corresponding sagittal and coronal T2-weighted turbo spin echo images of the spine (cf. below) by TH (>20 y of experience)]. The employed rAMIRA acquisition protocol was identical to the "optimized standard protocol version" presented in the corresponding methods paper (8). Thus, the relevant sequence parameters for rAMIRA were: field of view = 128 × 128 mm^2, 512 readout samples (includes 2 × oversampling), 260 projections, isotropic in-plane resolution 0.50 × 0.50 mm^2, slice thickness 8 mm, bandwidth = 310 Hz/Px, flip angle = 50 deg, signal averaging = 2. Five images with the mean inversion times TI_{eff} = 174, 239, 304, 368, 433 ms were acquired simultaneously and later combined (cf. below and **Figure 1**). The sequence uses cardiac triggering to mitigate potential pulsation artifacts, which was realized with a standard infrared finger clip (simple pulse triggering). Hence, for a heart rate of 60 bpm, rAMIRA's acquisition time corresponds to 2:39 min per slice, for instance.

Furthermore, all participants received T2-weighted turbo spin echo imaging covering the whole SC in sagittal and coronal slice orientation. Here, the most relevant sequence parameters were [1] sagittal: in-plane resolution = 0.7 × 0.7 mm^2, 17 slices of thickness 3 mm, TR = 3,400 ms, TE = 102 ms; [2] coronal: in-plane resolution = 1.4 × 1.4 mm^2, 17 slices of thickness 3 mm, TR = 3,500 ms, TE = 95 ms. Additionally, 3D isotropic high-resolution whole-brain T1-weighted images were acquired with the magnetization prepared rapid gradient echo (MPRAGE) sequence, using the following parameters: 1.0 mm isotropic resolution, TI = 1,100 ms, TR = 2,000 ms, TE = 2.12 ms, flip angle = 8 deg, matrix = 256 × 256 × 192.

MRI Analysis

All rAMIRA images were visually inspected for image quality [by TH (>20 y of experience), EK (1 y), MJW (3 y), and CW (>20 y)]. Only images with sufficient quality were segmented. Segmentations were not possible in 3 out of 244 images in the cervical SC and in 12 out of 122 in the thoracic SC (for details s. results section). For the quantitation of the GM area and TCA, one mean image of all five simultaneously acquired inversion images of the rAMIRA series was calculated, which shows a high gray matter to white matter contrast, as well as sufficient contrast between SC and CSF (**Figure 1**) (8, 9).

Total cross-sectional areas were segmented in a semi-automated way using the software JIM 7 (http:/www.xinapse.com) (31). Following a previously published reliable segmentation algorithm (5, 6, 26), GM was segmented manually three times by one single rater and the mean was calculated (by MJW).

Brain T1-weighted images were investigated using SPM12 (https://www.fil.ion.ucl.ac.uk/spm/) to determine the total intracranial volume (TIV; a frequently used normalization parameter for brain volumes) (32) as well as brain GM and WM volumes.

In addition, the following parameters were determined as potential normalization factors at the levels C4/C5 and Tmax (**Figure 2**) (by EK):

- Maximum axial anterior–posterior and lateral spinal canal diameters, axial spinal canal area, maximum axial lateral vertebral body width on the axial two-dimensional rAMIRA images; additionally, the product of the anterior–posterior and lateral spinal canal diameter was calculated at both levels.
- Middle vertebra height C4 and T12 on the T2-weighted sagittal images.
- Anterior–posterior and lateral diameters, as well as area of the foramen magnum, and distance between Basion–Opisthion on an isotropic, high-resolution 3D whole-brain T1-weighted image. The product of the anterior–posterior and lateral diameter of the foramen magnum was calculated as well.

The inter-rater reliability of the manually assessed anatomical parameters was determined in 14 datasets by two independent operators (by EK and CW) and showed excellent reliability with intra-class correlation coefficients (ICCs, two-way random, absolute agreement) (33) >0.96, as listed in the supplement (**Supplementary Table 1**). Segmentation was done with OsirixLite (https://www.osirix-viewer.com/).

Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics for Macintosh, Version 25.0, Armonk, NY: IBM Corp., and JMP pro, Version 14. SAS Institute Inc., Cary, NC, 1989-2019.

- Inter-Individual Variability Assessments

To assess the inter-individual variability of SC areas, the respective deviation from the group mean was calculated for each subject as

$$\frac{|\text{measured area} - \text{mean area}|}{\text{mean area}} * 100[\%]$$

- Assessing the Effect of Age and Sex

Differences in rAMIRA-based SC areas between men and women were assessed using linear regression analysis co-varying for age.

Linear and quadratic fits for (a) TCA and (b) SC GM areas vs. age were assessed for the whole study group ($n = 61$), and the r^2 of the models was reported. For practicability reasons to facilitate

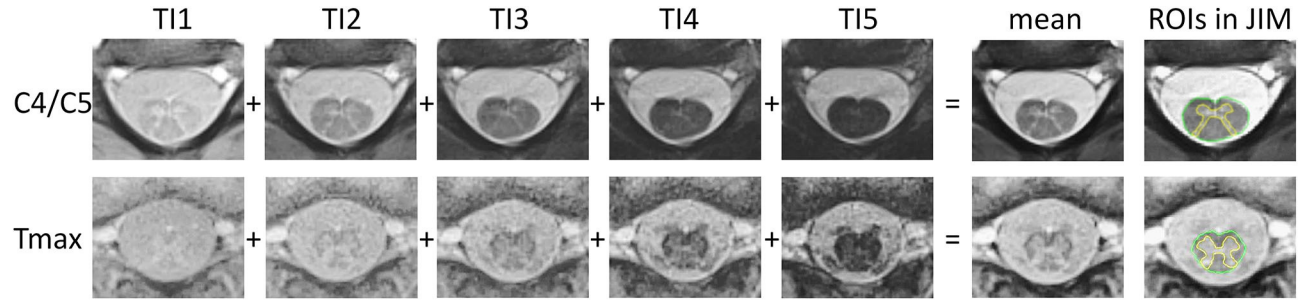


FIGURE 1 | rAMIRA images of the spinal cord in an 11-year-old boy. From left to right: axial rAMIRA images at intervertebral disc level C4/C5 (top row) and T_{max} (bottom row) acquired at five inversion times (TI), mean image (combination of the images at five TIs); and mean images with the region of interests (ROIs) spinal cord total cross-sectional area (TCA, green) and gray matter area (GMA, yellow) which were segmented using the software JIM.

future analyses in specific target populations, we also assessed SC areas in three nearly equally sized sub-groups: minors (Group 1: aged <18 years, $n = 18$), middle-aged (Group 2: aged 18–65 years, $n = 23$) and elderly subjects (Group 3: aged >65 years, $n = 20$) using linear regression analysis with sex as additional covariate, respectively.

- Normalization Models

For the development of the normalization models, the associations of the anatomical parameters with rAMIRA-based SC areas at C4/C5 and T_{max} were first assessed using Pearson correlation coefficients since all variables were normally distributed. Bonferroni correction was performed with a correction factor of 17 ($p < 0.05/17$) to correct for $n = 17$ tests. This analysis was performed using the data from all subjects ($n = 61$).

We then performed a backward selection procedure starting with a model containing all anatomical variables with a Pearson correlation coefficient of $|r| > 0.30$ (34) in univariate analysis as well as age and sex as predictor variables and TCA at the level C4/C5 as outcome parameter.

This procedure was performed (a) considering brain GM and WM volumes as potential predictors (approach suitable for studies in healthy controls and diseases known to not affect brain GM and/or WM volumes) and (b) without considering brain GM and WM volumes (approach suitable for studies in diseases known to affect these brain volumes).

The adjusted r^2 of the models resulting from the backward selection was reported.

For normalization we used the approach described by Sanfilipo and Papinutto (25, 35):

$$\mathrm{Area}_{\mathrm{predicted}} = \mathrm{Area}_{\mathrm{measured}} + \mathrm{a}(\mathrm{X}_{\mathrm{mean}} - \mathrm{X}_{\mathrm{measured}}) + \mathrm{b}(\mathrm{Y}_{\mathrm{mean}} - \mathrm{Y}_{\mathrm{measured}}) + \mathrm{c}(\mathrm{Z}_{\mathrm{mean}} - \mathrm{Z}_{\mathrm{measured}})$$

with a, b, and c being the estimates (regression coefficients) obtained by the linear regression analysis for the predictor variables surviving the backward selection procedure and X, Y, Z the measured values of these variables.

The performance of the resulting normalization models was expressed a) as the % reduction of inter-individual variability of the normalized SC areas of each model in relation to the variability of the non-normalized areas and b) as the % relative standard deviation (% RSD) reductions of the predicted areas to the % RSD of the non-normalized, measured areas. The relative standard deviation (RSD) is the standard deviation divided by the mean SC area.

The normalization model with the largest relative inter-individual variability reduction was then applied to all other SC level measurements.

In analogy, we developed a normalization model for the sub-group of minors ($n = 18$).

RESULTS

The acquired rAMIRA images displayed a good quality in general. Segmentations were not possible in 1.2% of the images in the cervical SC (due to motion artifacts in minors) and in 9.8% in the thoracic SC due to image artifacts originating from flow in near-by large vessels or due to artifacts originating from cardiac and breathing motion.

Inter-individual Variability of Spinal Cord Area Measurements in Healthy Subjects

The inter-individual mean relative variability for TCA and GM area, as well as the % relative SD (% RSD: defined as the SD divided by the group mean), are summarized in **Table 1**.

Effects of Age and Sex

Effects of age and sex are summarized in the supplement in **Supplementary Tables 2, 3**.

In brief, men showed significantly larger SC GM areas at C3/C4 ($p = 0.0350$) and at T_{max} ($p = 0.0497$) than women. At all other levels, we detected no significant area differences between sexes (**Supplementary Table 2**).

The linear and quadratic fits for TCA and GM areas vs. age for the whole group showed in general relatively low r^2 indicating low accuracy of the models: r^2 (linear/quadratic fit) $TCA_{C2/C3}$: 0.009/0.088; $TCA_{C3/C4}$: 0.005/0.049; $TCA_{C4/C5}$: 0.001/0.031; $TCA_{C5/C6}$: 0.010/0.034; $TCA_{T9/T10}$: 0.041/0.101; TCA_{Tmax}: 0.000/0.014; GM $area_{C2/C3}$: 0.051/0.055; GM $area_{C3/C4}$: 0.000/0.017; GM $area_{C4/C5}$: 0.009/0.018; GM

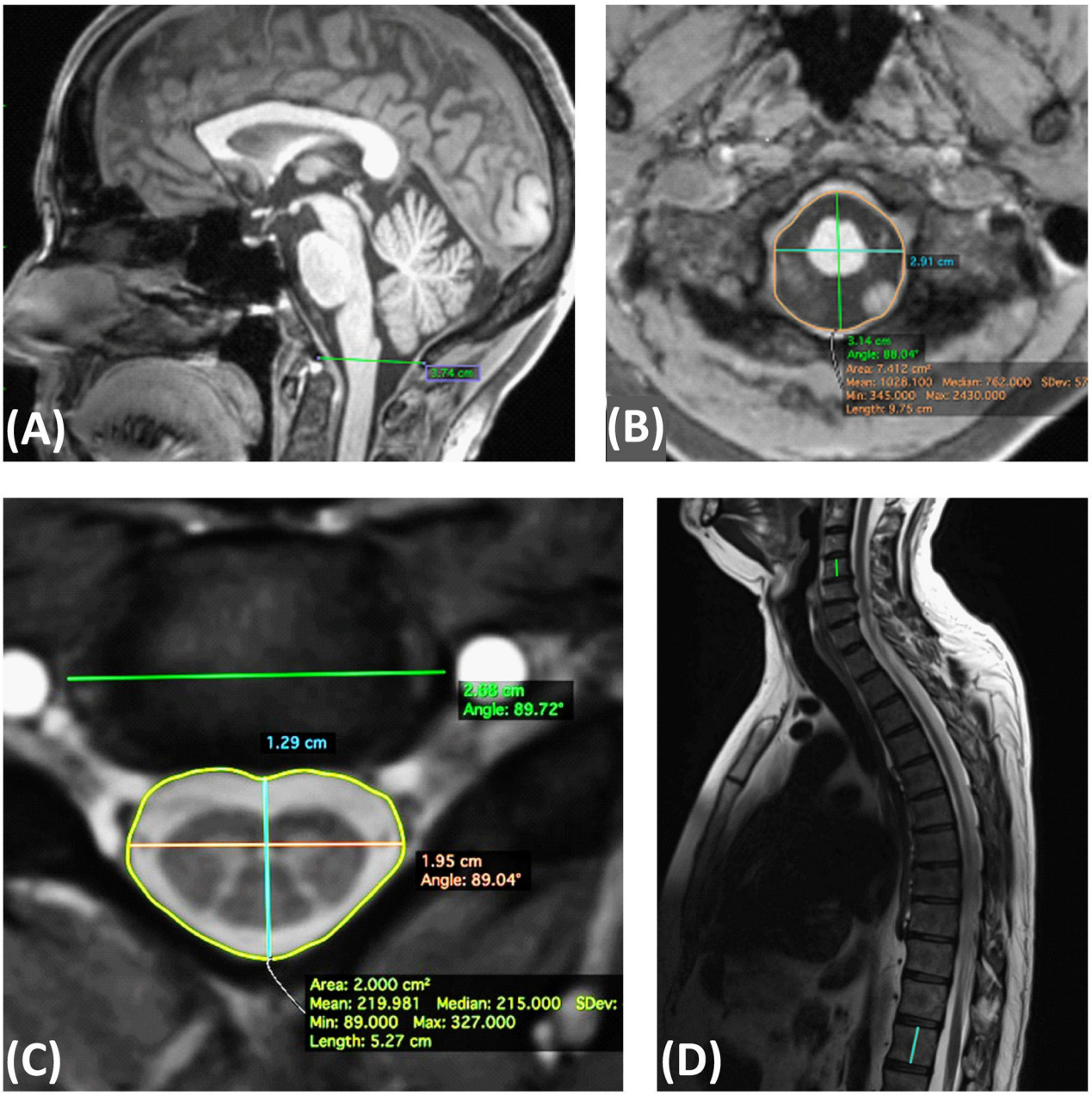


FIGURE 2 | Definition of the skeletal metrics. **(A)** Sagittal T1 weighted MPRAGE: McRae line (distance between Basion–Opisthion). **(B)** Axial MPRAGE: anterior posterior and lateral diameters, as well as area of the foramen magnum. **(C)** Axial rAMIRA: Maximum axial anterior posterior and lateral spinal canal diameters, spinal canal area, maximum lateral vertebral body width. **(D)** Sagittal T2 weighted turbo spin echo: middle vertebra height C4 and T12.

$area_{C5/C6}$: 0.038/0.045; and GM $area_{T9/T10}$: 0.229/0.24, GM $area_{Tmax}$: 0.13/0.137.

For reasons of practicability, we also analyzed SC areas in age sub-groups of minors, middle-aged, and elderly subjects with adjustment for sex. These results are summarized in **Supplementary Table 3**. Maximum mean TCA values were consistently observed in the sub-group of middle-aged subjects at all levels, with TCA differences being significant between minors and middle-aged subjects at the level C2/C3 (mean difference 4.78 mm^2, SE 2.34, 95% CI of the difference: 0.09–9.47, $p = 0.0457$), and between middle-aged and elderly subjects at the level T9/T10 (mean difference 3.30 mm^2, SE 1.28, 95% CI of the difference: 0.72–5.88, $p = 0.0131$).

While minimum mean GM area values were consistently observed in the subgroup of elderly subjects at all levels, maximum mean GM area values were observed in the

TABLE 1 | Mean % inter-individual variability, mean, SD and % RSD (% relative SD) of spinal cord total cross-sectional areas (TCA) and gray matter areas (GMA) at the intervertebral disc levels C2/C3–C5/C6, T9/T10, and T_{max} (level of the lumbar enlargement) of the whole study population.

Level		Mean (in mm^2)	SD	% RSD	Mean % inter-individual variability
C2/C3	TCA	83.2	7.5	9.0	6.6
	GMA	15.5	1.4	8.9	7.0
C3/C4	TCA	87.4	8.31	9.5	7.1
	GMA	19.1	2.0	10.4	8.2
C4/C5	TCA	88.5	9.0	10.1	8.1
	GMA	20.2	2.2	10.7	8.6
C5/C6	TCA	85.2	8.0	9.3	7.4
	GMA	20.0	2.2	11.1	9.0
T9/T10	TCA	46.8	4.4	9.3	7.3
	GMA	9.80	1.1	11.4	8.9
T_{max}	TCA	62.9	6.5	10.4	8.9
	GMA	24.0	3.8	15.6	12.2

subgroup of middle-aged subjects only at the level C3/C4 (**Supplementary Table 3**). Significant GM area differences could be detected between middle-aged and elderly subjects at the level C2/C3 (mean difference 0.83 mm^2, SE 0.41, 95% CI of the difference: 0.00–1.66, $p = 0.0495$) and between minors and middle-aged subjects at the level T9/T10 (mean difference 0.80 mm^2, SE 0.37, 95% CI: 0.06–1.55, $p = 0.0353$) (**Supplementary Table 3**).

Normalization Models

Among the chosen metrics, TIV, brain WM volume, anterior–posterior spinal canal diameter, and area of the spinal canal, as well as the product of the anterior–posterior and lateral spinal canal showed a correlation with TCA at the level C4/C5 with Pearson correlation coefficients of $|r| > 0.30$ (**Table 2**), with the other parameters showing lower correlation coefficients. After Bonferroni correction, none of the investigated metrics did show a significant association with SC areas at T_{max} (**Supplementary Table 4**); thus, normalization model development was primarily focused on the cervical SC.

The backward selection procedure yielded a model with sex, brain WM volume, and the area of the spinal canal at the level C4/C5 as predictor variables for TCA at C4/C5 as outcome (Model 1) (**Table 3**). Age was not a predictor variable.

Normalization strategies that are based on brain WM or GM volumes can be used in healthy subjects, but are not well-suited in diseases with potential changes in WM or GM volumes caused by the underlying disease itself (e.g., gray matter pathology in ALS, SMA; or mainly white matter pathology in adrenomyeloneuropathy).

We, therefore, performed a second backward selection—without considering brain GM and WM volumes—yielding a model (in analogy to Model 1) with sex, TIV, and the area of the spinal canal at the level C4/C5 as predictors for TCA at C4/C5 as outcome, permitting variables with $p < 0.1$ (Model 1a). This model was further simplified using backward selection to a model only containing the area of the spinal canal at the level C4/C5 as univariate predictor (Model 2) (**Table 3**).

The effect of the normalization of TCA and GM areas at the level C4/C5 based on Models 1, 1a, and 2 on the mean inter-individual variability and % relative standard deviation is summarized in **Table 4**.

Model 1 and Model 2 were then applied to all SC level measurements (**Table 5**, **Figure 3**).

TABLE 2 | Associations between the anatomical metrics and total cross-sectional area (TCA) and gray matter area (GMA) at the C4/C5 level using Pearson correlation coefficients.

Whole study population ($n = 61$)	TCA C4/C5		GMA C4/C5	
Metric	***p*-Value**	**Pearson correlation coefficient**	***p*-Value**	**Pearson correlation coefficient**
CAN_C4/C5_ap	**0.000**	**0.43**	**0.000**	**0.47**
CAN_C4/C5_lat	0.470	0.26	0.178	0.18
CAN_C4/C5_area	**0.000**	**0.44**	**0.001**	**0.43**
Prod_CAN_C4/C5_ap*lat	**0.000**	**0.45**	**0.000**	**0.44**
VBW_C4/C5_lat	0.706	0.05	0.315	−0.13
VBH_C4	0.519	0.08	0.860	0.02
McRae	0.532	0.08	0.760	0.04
ForMag_ap	0.943	0.01	0.821	0.03
ForMag_lat	0.998	0.00	0.493	0.09
ForMag_area	0.377	0.12	0.201	0.17
Prod_ForMag_ap*lat	0.992	−0.00	0.892	−0.02
TIV	0.012	**0.32**	0.074	0.23
Brain GM volume	0.394	0.11	0.175	0.18
Brain WM volume	**0.000**	**0.48**	0.048	0.26
Height	0.197	0.17	0.650	0.06
Weight	0.170	0.19	0.211	0.16
BMI	0.296	0.14	0.155	0.18

*P-values surviving the Bonferroni correction are bolded ($p < 0.0029$). CAN_C4/C5_ap, anterior–posterior diameter of the spinal canal at the level C4/C5. CAN_C4/C5_lat, lateral diameter of the spinal canal at the level C4/C5. CAN_C4/C5_area, area of the spinal canal at the level C4/C5; Prod_CAN_C4/C5_ap*lat, product of the anterior-posterior and lateral diameter of the spinal canal at the level C4/C5; VBW_C4/C5_lat, maximum vertebra body width at the level C4/C5; VBH_C4, middle vertebral body height of C4; McRae, McRae line (distance between Basion–Opisthion). ForMag_ap, anterior posterior diameter of the foramen magnum; ForMag_lat, lateral diameter of the foramen magnum; ForMag_area, area of the foramen magnum; Prod_ForMag_ap*lat, product of the anterior–posterior and lateral diameter of the foramen magnum; TIV, total intracranial volume; GM, gray matter; WM, white matter; BMI, body mass index.*

Normalization Models for Minors

In a subgroup analysis, we separately examined potential normalization variables for anatomical variability reduction in minors ($n = 18$) at the C4/C5 level. **Supplementary Table 5** in the supplement shows the univariate associations between SC areas and potential covariates using Pearson correlation.

Backward selection (starting with a model containing all variables with a Pearson correlation coefficient of $|r| > 0.30$

TABLE 3 | Linear regression analysis with total cross-sectional area (TCA) at the C4/C5 level as outcome and normalization variables after backward selection; corresponding models with gray matter area (GMA) as outcome.

Whole study population	TCA C4/C5			GMA C4/C5		
	p	Adjusted r^2	Estimate	*p*	Adjusted r^2	Estimate
Model 1:	<0.0001	0.43		0.0014	0.20	
Sex	0.0082		−2.8534	0.4019		−0.2523
WM volume	<0.0001		0.0861	0.0495		0.0096
CAN_C4C5_area	<0.0001		0.1235	0.0008		0.0266
Model 1a:	0.0005	0.23		0.0057	0.15	
Sex	0.0975		−2.1357	0.5926		−0.1727
TIV	0.0271		0.0177	s 0.3237		0.0020
CAN_C4C5_area	0.0010		0.1063	0.0024		0.0247
Model 2:	0.0004	0.18		0.0006	0.17	
CAN_C4C5_area	0.0004		0.1094	0.0006		0.0257

CAN_C4C5_area, spinal canal area at the level C4/C5; TIV, total intracranial volume.

TABLE 4 | % relative standard deviation (RSD, standard deviation divided by the mean area), relative % RSD reduction, mean % inter-individual variability [(measured area in a given subject–group mean area)/group mean area], and relative % inter-individual variability reduction with respect to the measured total cross-sectional cord area (TCA) and gray matter area (GMA) at the intervertebral disc level C4/C5 for normalizations based on Models 1, 1a, and Model 2.

Area		Non-normalized	Model 1	Model 1a	Model 2
TCA C4/C5	% RSD	10.1	7.7	8.8	9.1
	Relative % RSD reduction [%]	–	(23.7)	(12.8)	(10.2)
	Mean % inter-individual variability	8.1	5.9	7.0	7.3
	Relative %variability reduction [%]		(27.2)	(12.9)	(10.0)
GMA C4/C5	% RSD	10.7	9.4	9.6	9.7
	Relative % RSD reduction [%]	–	(12.0)	(10.0)	(9.6)
	Mean % inter-individual variability	8.6	7.5	7.9	7.9
	Relative %variability reduction [%]	–	(13.2)	(8.1)	(8.6)

Whole study population, n = 61.

TABLE 5 | Performance of normalization by Model 1 and Model 2 for total cross-sectional cord areas (TCA) and gray matter areas (GMA) at all other cord levels.

Level		% RSD measured	Model 1	Model 2
			% RSD normalized [relative % RSD reduction (%)]	% RSD normalized [relative % RSD reduction (%)]
C2/C3	TCA	9.0	7.3 (19.0)	8.9 (1.4)
	GMA	8.9	7.9 (11.3)	8.3 (6.9)
C3/C4	TCA	9.5	7.9 (17.1)	9.3 (1.9)
	GMA	10.4	9.5 (9.2)	10.2 (2.0)
C4/C5	TCA	10.1	7.7 (23.7)	9.1 (10.2)
	GMA	10.7	9.4 (12.0)	9.7 (9.6)
C5/C6	TCA	9.3	7.2 (23.5)	8.9 (4.5)
	GMA	11.1	10.0 (9.8)	10.7 (3.6)
T9/T10	TCA	9.3	8.0 (13.9)	9.1 (2.7)
	GMA	11.4	10.9 (4.1)	11.1 (3.0)
T_{max}	TCA	10.4	9.6 (7.5)	10.3 (1.3)
	GMA	15.6	14.9 (4.6)	15.5 (0.8)

% relative standard deviation (RSD) and relative % RSD reduction obtained by normalization based on Model 1 (sex, brain WM volume and spinal canal area at the level C4/C5) as well as by Model 2 (spinal canal area at the level C4/C5) (whole study population, n = 61).

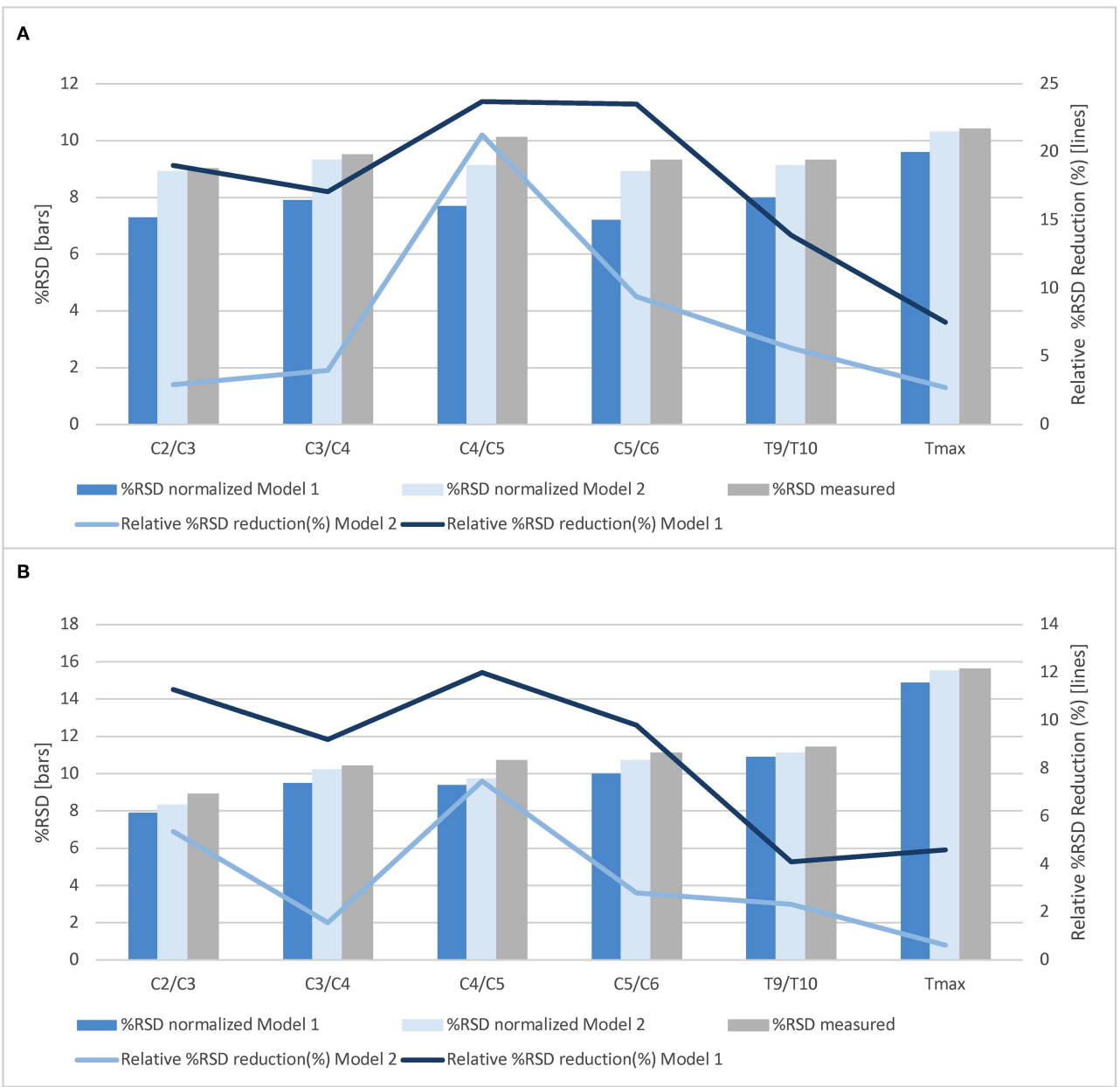


FIGURE 3 | Normalization effects of Model 1 and Model 2 on total cross-sectional cord areas and gray matter areas at all cord levels. The % relative standard deviation (RSD, standard deviation divided by the mean area) and relative % RSD reduction obtained by normalization based on Model 1 (sex, brain white matter volume and spinal canal area at the level C4/C5) as well as by Model 2 (spinal canal area at the level C4/C5) for **(A)** total cross-sectional cord areas and **(B)** spinal cord gray matter areas (whole study population, $n = 61$).

in univariate analysis together with age and sex as predictor variables and TCA at C4/C5 as outcome parameter) resulted in a model with the area of the spinal canal at level C4/C5 and weight as predictor variables (Model 3 m) (p of the model = 0.0027, adjusted $r^2 = 0.48$). Normalization reduced the RSD about 32.5% for TCA and 18.3% for GM area at the level C4/C5. Inter-individual variability was reduced from 8.1 to 5.7% for TCA and reduced from 9.3 to 7.1% for GM area. This normalization model was applied to all other measured SC levels (**Table 6**).

For reasons of practicability, Model 3 m was further simplified to a univariate model containing the area of the spinal canal of the C4/C5 level as single predictor variable in minors (in analogy to Model 2 named Model 2 m) ($p = 0.0074$, adjusted $r^2 = 0.33$). This normalization model was also applied to all other measured SC levels in minors (**Table 6**).

TABLE 6 | Performance of normalization by Model 3 m and Model 2 m for total cross-sectional cord areas (TCA) and gray matter areas (GMA) at all other cord levels in minors.

Minors level		% RSD measured	Model 3 m	Model 2 m
			% RSD normalized [relative % RSD reduction (%)]	% RSD normalized [relative % RSD reduction (%)]
C2/C3	TCA	9.1	7.0 (23.1)	7.5 (17.6)
	GMA	10.1	6.8 (32.8)	6.9 (31.9)
C3/C4	TCA	9.8	6.7 (31.8)	7.9 (18.9)
	GMA	10.9	8.7 (20.7)	9.3 (15.4)
C4/C5	TCA	11.3	7.6 (32.5)	9.0 (20.6)
	GMA	11.4	9.3 (18.3)	9.7 (14.5)
C5/C6	TCA	9.3	5.0 (45.9)	6.1 (34.5)
	GMA	11.6	7.6 (34.4)	8.5 (26.3)
T9/T10	TCA	9.0	6.2 (31.9)	8.0 (12.0)
	GMA	6.5	5.1 (20.3)	5.3 (17.2)
T_{max}	TCA	9.6	8.8 (9.2)	9.6 (0.4)
	GMA	10.1	8.2 (19.3)	8.5 (16.3)

% relative standard deviation (RSD) and relative % RSD reduction of normalization by Model 3 m (spinal canal area at the C4/C5 level and weight) and Model 2 m (spinal canal area at the C4/C5 level) applied to all other measured cord levels in minors.

DISCUSSION

Anatomical inter-subject variations in healthy subjects are a relevant source of SC TCA and GM area variability. Our study demonstrated an inter-individual variability of TCA and SC GM areas ranging from 6 to 9% in the cervical and 9–12% in the thoracic SC, which is in line with prior reports investigating the upper cervical SC (24).

Before applying morphometric SC GM/WM imaging techniques such as the novel rAMIRA approach for atrophy assessments in clinical studies in lower motor neuron disorders, it therefore seems necessary to develop efficient strategies to reduce the inter-subject variability of SC area measurements in healthy subjects to facilitate the detection of pathology-related changes.

This study explored potential normalization strategies for variability reduction in 61 healthy subjects with a broad age range from 11 to 93 years (including a subgroup of 18 minors).

Normalization models were developed using backward selection starting with a model containing the predictor variables age and sex as well as all anatomical metrics with a Pearson correlation coefficient of $|r| > 0.3$ in univariate analysis with TCA at C4/C5—the level of the cervical enlargement—as outcome. The backward selection procedure yielded a model including the predictor variables sex, spinal canal area at C4/C5, and brain WM volume. Normalization of SC areas by this model reduced the % RSD of TCA at the level C4/C5 by 24%, of GM area by 12%, and also at the other cervical SC levels in the range of 17–24% (TCA) and 9–12% (GM area). Therefore, this approach seems an effective normalization strategy suited for studies in healthy subjects or in SC diseases that do not affect cerebral WM.

However, as brain WM volume is frequently altered in SC diseases either directly by a disease pathology affecting both brain and SC (e.g., in multiple sclerosis) or potentially also secondary to lower motor neuron disorders, e.g., by retrograde trans-synaptic degeneration, we also performed a second backward analysis containing only demographic and skeleton-based variables otherwise following the above mentioned selection criteria. The resulting model contained the spinal canal area as single predictor. Normalization by this model reduced the % RSD of TCA and GM area at C4/C5 by 10%, respectively. However, application of this normalization to other SC levels, particularly to the thoracic SC, showed only small effects in our cohort, with relative % RSD reductions ranging from 1 to 7%.

Age was not a significant predictor variable for SC areas in this study. During normalization model development, age consistently was eliminated in the backward selection process. We consistently observed maximum mean TCA values in the middle-aged group at all SC levels. This is in line with the observation by Papinutto et al. (24) of a TCA peak at the level C2/C3 at ∼45 years of age. While at all levels minimum mean SC GM area values were observed in the subgroup of elderly subjects, maximum mean SC GM area values were observed in the subgroup of middle-aged subjects only at the level C3/C4, with equally high or higher mean SC GM area values in minors at all other levels.

Since children and adolescents are the main target group for the recently approved SMA treatments (36) and no data exists on SC GM area inter-subject variability, we conducted a subgroup analysis in minors ($n = 18$). Following the same variable selection procedure as described above, normalization based on the spinal canal area at the level C4/C5 and the variable body weight resulted in a reduction of % RSD of TCA values at the level C4/C5 by 32%. Analogous normalization of GM areas reduced the % RSD by 18% at the level C4/C5, and also consistently at

all other cervical and thoracic levels by 19–34%. The % RSD reduction observed in TCAs by this normalization method are slightly larger than what has been described in a very recently published study in minors aged 7–17 years by normalization with the product of the anterior posterior and lateral diameter of the spinal canal at the level C2/3 and skull volume (30). This difference could be partly explained by difference in in-plane resolution of the images used for SC segmentation (0.5 vs. 1 mm^2) (30) and by the chosen predictors. Effects on SC GM area were not investigated in that study (30).

The spinal canal area is relatively easy to measure and showed excellent inter-rater reliability coefficients (33). Normalization based on spinal canal area at the level C4/C5 and body weight could therefore be a promising and brain volume-independent procedure in upcoming studies involving minors with lower motor neuron diseases to reduce anatomical inter-subject variability. Whether this approach increases the sensitivity and specificity in detecting disease-related changes in motor neuron disorders or treatment effects, needs to be further investigated in subsequent studies involving patients.

Image quality of the cervical rAMIRA images was generally high permitting SC, TCA, and GM segmentation in >98% of acquired cervical SC images of the study population. Imaging the thoracic SC is in general more challenging: Factors, that can negatively impact image quality at the thoracic levels include lower coil sensitivity in this area (**Figure 1**, bottom row) and potential artifacts arising from heart and breathing motion, from pulsating large blood vessels, and from increased susceptibility variations. Despite these limitations, the novel rAMIRA method enabled TCA and GM area segmentation in >90% of the acquired thoracic SC images. Segmentation was not reliably possible at the level T9/T10 and T_{max} in 6 out of 61 subjects, respectively, because of the upper mentioned issues. The % RSD of the SC metrics was higher at T_{max} compared to the cervical levels (**Table 1**), which is likely due to the high inherent anatomical variability in this region. None of the investigated regional anatomical metrics did show a significant association with SC areas at $T_{max.}$ A unifying normalization concept based on normalization of SC areas at different levels, by one model based on anatomical metrics measured at one selected level, and not by several models with regional normalization metrics, was therefore chosen. Though not perfect, this approach will also permit comparison of atrophy effects between levels.

It should be mentioned that the study has the following limitations: the sample size of 61 healthy subjects, including only 18 minors, is rather small. The assumption that the investigated skeleton-derived metrics are fully disease independent might be questioned, as secondary orthopedic complications, immobilization, and also treatment effects might alter spine metrics in a way that is not easily predicted. Spine-derived metrics were assessed manually in this study but showed excellent inter-rater reliability.

In conclusion, rAMIRA imaging is a novel approach to perform SC imaging with favorable contrast in clinical settings that is well-suited for GM/WM quantitation in both the cervical and thoracic SC. Our study demonstrates that the inter-individual variability of SC area measurements in healthy subjects is relatively high but can be effectively reduced in the cervical and to a lesser degree also in the thoracic SC, particularly in minors, by a model based on spinal canal area at the level C4/C5 and weight. By reducing anatomical inter-subject variability, these approaches may facilitate the detection of pathology-related changes and the improvement of therapeutic monitoring in lower motor neuron disorders in the future.

DATA AVAILABILITY STATEMENT

Upon reasonable request, we will render the detailed results derived from the reported analyses available.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission Nordwest- und Zentralschweiz - EKNZ. Written informed consent to participate in this study was provided by the participants or the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

EK, MW, CW, OB, and RS conceptualized and designed the study. MW, CW, AH, PC, and OB handled the methods development. EK, JW, LR, TH, and RS collected the data. EK, JW, CW, TH, MB, LS, and RS conducted the data analysis. EK, MW, CW, and RS were in charge of the paper drafting. EK, JW, MW, CW, TH, LR, LS, AH, MB, PC, CG, OB, and RS reviewed the manuscript. OB and RS obtained the funding. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Swiss National Science Foundation (SNSF) (Grant No 320030_156860); Stiftung zur Förderung der gastroenterologischen und allgemeinen klinischen Forschung sowie der medizinischen Bildauswertung, and Biogen Inc. The funder did not play any role in study design, data, analysis and interpretation of the data.

ACKNOWLEDGMENTS

The authors thank all study participants.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.637198/full#supplementary-material

REFERENCES

1. Fradet L, Arnoux PJ, Ranjeva JP, Petit Y, Callot V. Morphometrics of the entire human spinal cord and spinal canal measured from *in vivo* high-resolution anatomical magnetic resonance imaging. *Spine.* (2014) 39:E262–9. doi: 10.1097/BRS.0000000000000125
2. Taso M, Le Troter A, Sdika M, Ranjeva JP, Guye M, Bernard M, et al. Construction of an *in vivo* human spinal cord atlas based on high-resolution MR images at cervical and thoracic levels: preliminary results. *MAGMA*. (2014) 27:257–67. doi: 10.1007/s10334-013-0403-6
3. Yiannakas MC, Kakar P, Hoy LR, Miller DH, Wheeler-Kingshott CA. The use of the lumbosacral enlargement as an intrinsic imaging biomarker: feasibility of grey matter and white matter cross-sectional area measurements using MRI at 3T. *PLoS ONE.* (2014) 9:e105544. doi: 10.1371/journal.pone.0105544
4. Yiannakas MC, Kearney H, Samson RS, Chard DT, Ciccarelli O, Miller DH, et al. Feasibility of grey matter and white matter segmentation of the upper cervical cord *in vivo*: a pilot study with application to magnetization transfer measurements. *Neuroimage.* (2012) 63:1054–9. doi: 10.1016/j.neuroimage.2012.07.048
5. Kearney H, Yiannakas MC, Abdel-Aziz K, Wheeler-Kingshott CA, Altmann DR, Ciccarelli O, et al. Improved MRI quantification of spinal cord atrophy in multiple sclerosis: cord atrophy in MS. *J Magn Reson Imaging.* (2014) 39:617–23. doi: 10.1002/jmri.24194
6. Papinutto N, Schlaeger R, Panara V, Caverzasi E, Ahn S, Johnson KJ, et al. 2D phase-sensitive inversion recovery imaging to measure *in vivo* spinal cord gray and white matter areas in clinically feasible acquisition times. *J Magn Reson Imaging.* (2015) 42:698–708. doi: 10.1002/jmri.24819
7. Tsagkas C, Horvath A, Altermatt A, Pezold S, Weigel M, Haas T, et al. Automatic spinal cord gray matter quantification: a novel approach. *AJNR Am J Neuroradiol.* (2019) 40:1592–600. doi: 10.3174/ajnr.A6157
8. Weigel M, Haas T, Wendebourg MJ, Schlaeger R, Bieri O. Imaging of the thoracic spinal cord using radially sampled averaged magnetization inversion recovery acquisitions. *J Neurosci Methods.* (2020) 343:108825. doi: 10.1016/j.jneumeth.2020.108825
9. Weigel M, Bieri O. Spinal cord imaging using averaged magnetization inversion recovery acquisitions. *Magn Reson Med.* (2018) 79:1870–81. doi: 10.1002/mrm.26833
10. Sabaghian S, Dehghani H, Batouli SAH, Khatibi A, Oghabian MA. Fully automatic 3D segmentation of the thoracolumbar spinal cord and the vertebral canal from T2-weighted MRI using K-means clustering algorithm. *Spinal Cord.* (2020) 58:811–20. doi: 10.1038/s41393-020-0429-3
11. Gros C, De Leener B, Badji A, Maranzano J, Eden D, Dupont SM, et al. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. *Neuroimage.* (2019) 184:901–15. doi: 10.1016/j.neuroimage.2018.09.081
12. Prados F, Cardoso MJ, Yiannakas MC, Hoy LR, Tebaldi E, Kearney H, et al. Fully automated grey and white matter spinal cord segmentation. *Sci Rep.* (2016) 6:36151. doi: 10.1038/srep36151
13. Datta E, Papinutto N, Schlaeger R, Zhu A, Carballido-Gamio J, Henry RG. Gray matter segmentation of the spinal cord with active contours in MR images. *Neuroimage.* (2017) 147:788–99. doi: 10.1016/j.neuroimage.2016.07.062
14. Horvath A, Tsagkas C, Andermatt. S, Pezold S, Parmar K, Cattin P. *Spinal Cord Gray Matter-White Matter Segmentation on Magnetic Resonance AMIRA Images With MD-GRU.* (2018). Available online at: https://arxiv.org/pdf/1808.02408.pdf (accessed March 04, 2021).
15. Blaiotta C, Freund P, Curt A, Cardoso J, Ashburner J. A probabilistic framework to learn average shaped tissue templates and its application to spinal cord image segmentation. In: *Proceedings of the 24th Annual Meeting of ISMRM*. Singapore (2016). p. 1449.
16. Porisky A, Brosch T, Ljungberg E,Tang LYW, Yoo Y, De Leener B, Traboulsee A, et al. Grey matter segmentation in spinal cord MRIs via 3D convolutional encoder networks with shortcut connections. In: Cardoso MJ, Arbel T, Carneiro G, editors. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer International Publishing (2017). p. 330–7.
17. Dupont SM, De Leener B, Taso M, Le Troter A, Nadeau S, Stikov N, et al. Fully-integrated framework for the segmentation and registration of the spinal cord white and gray matter. *Neuroimage.* (2017) 150:358–72. doi: 10.1016/j.neuroimage.2016.09.026
18. Prados F, Ashburner J, Blaiotta C, Brosch T, Carballido-Gamio J, Cardoso MJ, et al. Spinal cord grey matter segmentation challenge. *Neuroimage.* (2017) 152:312–29. doi: 10.1016/j.neuroimage.2017.03.010
19. Engl C, Schmidt P, Arsic M, Boucard CC, Biberacher V, Röttinger M, et al. Brain size and white matter content of cerebrospinal tracts determine the upper cervical cord area: evidence from structural brain MRI. *Neuroradiology.* (2013) 55:963–70. doi: 10.1007/s00234-013-1204-3
20. Mann RS, Constantinescu CS, Tench CR. Upper cervical spinal cord cross-sectional area in relapsing remitting multiple sclerosis: application of a new technique for measuring cross-sectional area on magnetic resonance images. *J Magn Reson Imaging.* (2007) 26:61–5. doi: 10.1002/jmri.20959
21. Song F, Huan Y, Yin H, Ge Y, Wei G, Chang Y, et al. Normalized upper cervical spinal cord atrophy in multiple sclerosis. *J Neuroimaging.* (2008) 18:320–7. doi: 10.1111/j.1552-6569.2007.00222.x
22. Healy BC, Arora A, Hayden DL, Ceccarelli A, Tauhid SS, Neema M, et al. Approaches to normalization of spinal cord volume: application to multiple sclerosis. *J Neuroimaging.* (2012) 22:e12–9. doi: 10.1111/j.1552-6569.2011.00629.x
23. Oh J, Seigo M, Saidha S, Sotirchos E, Zackowski K, Chen M, et al. Spinal cord normalization in multiple sclerosis. *J Neuroimaging.* (2014) 24:577–84. doi: 10.1111/jon.12097
24. Papinutto N, Asteggiano C, Bischof A, Gundel TJ, Caverzasi E, Stern WA, et al. Intersubject variability and normalization strategies for spinal cord total cross-sectional and gray matter areas. *J Neuroimaging.* (2020) 30:110–8. doi: 10.1111/jon.12666
25. Sanfilipo MP, Benedict RHB, Zivadinov R, Bakshi R. Correction for intracranial volume in analysis of whole brain atrophy in multiple sclerosis: the proportion vs. residual method. *Neuroimage.* (2004) 22:1732–43. doi: 10.1016/j.neuroimage.2004.03.037
26. Schlaeger R, Papinutto N, Panara V, Bevan C, Lobach IV, Bucci M, et al. Spinal cord gray matter atrophy correlates with multiple sclerosis disability. *Ann Neurol.* (2014) 76:568–80. doi: 10.1002/ana.24241
27. Casserly C, Seyman EE, Alcaide-Leon P, Guenette M, Lyons C, Sankar S, et al. Spinal cord atrophy in multiple sclerosis: a systematic review and meta-analysis: spinal cord atrophy in multiple sclerosis: systematic review. *J Neuroimaging.* (2018) 28:556–86. doi: 10.1111/jon.12553
28. Isik S, Samanci MY. Analysis of MRI morphometric parameters of the pediatric cervical spine and spinal cord. *J Turk Spinal Surg.* (2019) 30:157–61. Available online at: http://cms.galenos.com.tr/Uploads/Article_27965/jtss-30-157-En.pdf (accessed March 02, 2021).
29. Lao L, Chen Z, Qiu G, Shen J. Whole-spine magnetic resonance imaging study in healthy Chinese adolescents. *Orthop Surg.* (2013) 5:164–70. doi: 10.1111/os.12046
30. Papinutto N, Cordano C, Asteggiano C, Caverzasi E, Mandelli ML, Lauricella M, et al. MRI measurement of upper cervical spinal cord cross-sectional area in children. *J Neuroimaging.* (2020) 30:598–602. doi: 10.1111/jon.12758
31. Horsfield MA, Sala S, Neema M, Absinta M, Bakshi A, Sormani MP, et al. Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis. *Neuroimage.* (2010) 50:446–55. doi: 10.1016/j.neuroimage.2009.12.121
32. Malone IB, Leung KK, Clegg S, Barnes J, Whitwell JL, Ashburner J, et al. Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. *NeuroImage.* (2015) 104:366–72. doi: 10.1016/j.neuroimage.2014.09.034
33. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* (2016) 15:155–63. doi: 10.1016/j.jcm.2016.02.012
34. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: L. Erlbaum Associates (1988).
35. Papinutto N, Schlaeger R, Panara V, Zhu AH, Caverzasi E, Stern WA, et al. Age, gender and normalization covariates for spinal cord gray matter and total cross-sectional areas at cervical and thoracic levels: a 2D phase sensitive inversion recovery imaging

study. *PLoS ONE.* (2015) 10:e0118576. doi: 10.1371/journal.pone.0118576

36. Waldrop MA, Kolb SJ. Current treatment options in neurology-SMA therapeutics. *Curr Treat Options Neurol.* (2019) 21:25. doi: 10.1007/s11940-019-0568-z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 29 March 2021
doi: 10.3389/fneur.2021.613834

Pilot Study on Quantitative Cervical Cord and Muscular MRI in Spinal Muscular Atrophy: Promising Biomarkers of Disease Evolution and Treatment?

Giovanni Savini[1], Carlo Asteggiano[1,2], Matteo Paoletti[1]*, Stefano Parravicini[2], Elena Pezzotti[2], Francesca Solazzo[1], Shaun I. Muzic[2], Francesco Santini[3,4], Xeni Deligianni[3,4], Alice Gardani[5], Giancarlo Germani[1], Lisa M. Farina[1], Niels Bergsland[6,7], Claudia A. M. Gandini Wheeler-Kingshott[2,8,9], Angela Berardinelli[5], Stefano Bastianello[1,2] and Anna Pichiecchio[1,2]*

[1] Advanced Imaging and Radiomics Center, Neuroradiology Department, IRCCS Mondino Foundation, Pavia, Italy, [2] Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy, [3] Department of Radiology, Division of Radiological Physics, University Hospital Basel, Basel, Switzerland, [4] Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland, [5] Child Neuropsychiatry Unit, IRCCS Mondino Foundation, Pavia, Italy, [6] Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, The State University of New York, Buffalo, NY, United States, [7] IRCCS, Fondazione Don Carlo Gnocchi ONLUS, Milan, Italy, [8] NMR Research Unit, Queen Square MS Centre, Department of Neuroinflammation, UCL Queen Square Institute of Neurology, University College London, Russell Square, London, United Kingdom, [9] Brain Connectivity Research Unit, IRCCS Mondino Foundation, Pavia, Italy

Introduction: Nusinersen is a recent promising therapy approved for the treatment of spinal muscular atrophy (SMA), a rare disease characterized by the degeneration of alpha motor neurons (αMN) in the spinal cord (SC) leading to progressive muscle atrophy and dysfunction. Muscle and cervical SC quantitative magnetic resonance imaging (qMRI) has never been used to monitor drug treatment in SMA. The aim of this pilot study is to investigate whether qMRI can provide useful biomarkers for monitoring treatment efficacy in SMA.

Methods: Three adult SMA 3a patients under treatment with nusinersen underwent longitudinal clinical and qMRI examinations every 4 months from baseline to 21-month follow-up. The qMRI protocol aimed to quantify thigh muscle fat fraction (FF) and water-T2 (w-T2) and to characterize SC volumes and microstructure. Eleven healthy controls underwent the same SC protocol (single time point). We evaluated clinical and imaging outcomes of SMA patients longitudinally and compared SC data between groups transversally.

Results: Patient motor function was stable, with only Patient 2 showing moderate improvements. Average muscle FF was already high at baseline (50%) and progressed over time (57%). w-T2 was also slightly higher than previously published data at baseline and slightly decreased over time. Cross-sectional area of the whole SC, gray matter (GM), and ventral horns (VHs) of Patients 1 and 3 were reduced compared to controls and remained stable over time, while GM and VHs areas of Patient 2 slightly increased.

OPEN ACCESS

Edited by:
Giuseppe Piscosquito, Fondazione Salvatore Maugeri (IRCCS), Italy

Reviewed by:
Elena Maria Pennisi, Ospedale San Filippo Neri, Italy
Valeria Ada Sansone, University Hospital Policlinico G. Martino, Italy

***Correspondence:**
Giovanni Savini
giovanni.savini@mondino.it
Matteo Paoletti
matteo.paoletti@mondino.it

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: 03 October 2020
Accepted: 15 February 2021
Published: 29 March 2021

Citation:
Savini G, Asteggiano C, Paoletti M, Parravicini S, Pezzotti E, Solazzo F, Muzic SI, Santini F, Deligianni X, Gardani A, Germani G, Farina LM, Bergsland N, Gandini Wheeler-Kingshott CAM, Berardinelli A, Bastianello S and Pichiecchio A (2021) Pilot Study on Quantitative Cervical Cord and Muscular MRI in Spinal Muscular Atrophy: Promising Biomarkers of Disease Evolution and Treatment? Front. Neurol. 12:613834. doi: 10.3389/fneur.2021.613834

We found altered diffusion and magnetization transfer parameters in SC structures of SMA patients compared to controls, thus suggesting changes in tissue microstructure and myelin content.

Conclusion: In this pilot study, we found a progression of FF in thigh muscles of SMA 3a patients during nusinersen therapy and a concurrent slight reduction of w-T2 over time. The SC qMRI analysis confirmed previous imaging and histopathological studies suggesting degeneration of αMN of the VHs, resulting in GM atrophy and demyelination. Our longitudinal data suggest that qMRI could represent a feasible technique for capturing microstructural changes induced by SMA *in vivo* and a candidate methodology for monitoring the effects of treatment, once replicated on a larger cohort.

Keywords: spinal muscular atrophy, muscle MRI, nusinersen (spinraza), motor neuron degeneration, spinal cord atrophy and degeneration, muscle fat fraction, quantitative magnetic resonance imaging, treatment monitoring

INTRODUCTION

Spinal muscular atrophy (SMA) refers to a spectrum of genetic neuromuscular disorders characterized by the degeneration of alpha motor neurons (αMNs) in the spinal cord (SC) (1) that leads to progressive muscle atrophy, consequent loss of muscular strength and paralysis. Within this group of disorders, the so-called 5q SMA is a rare condition (estimated incidence: from 1/6,000 to 1/10,000 live births) (2) caused by homozygous loss of function of SMN1 gene (survival motor neuron gene 1—telomeric form) (5q11.2-q11.3), with consequent SMN protein deficiency in lower motor neurons. SMA phenotypes encompass a continuum ranging from extremely severe neonatal onset to mild adult onset (3). Phenotypic variability in SMA is largely determined by the copy number of the SMN2 gene. Each SMN2 copy only codes for 10–15% of functional SMN proteins. In all SMA patients, because of the SMN1 deletion, SMN levels in neurons depend on SMN protein coded by the SMN2 gene, which in turn depends on the copy number of SMN2 gene. Therefore, the SMN2 copy number inversely correlates with disease severity (4, 5).

Once the role of the SMN protein expression from SMN2 was recognized, several strategies were developed to increase SMN protein expression (6). Three approved therapies for the treatment of SMA are currently available: nusinersen, a 2-O-methoxyethyl phosphorothioate-modified antisense oligonucleotide that modifies the splicing of the SMN2 pre-mRNA, increasing the production of full-length SMN protein from the SMN2 gene; onasemnogene abeparvovec (AVXS-101), a gene replacement therapy; and risdiplam, an orally administered small-molecule SMN2 splicing modifier that was recently approved by the Food and Drug Administration for the treatment of SMA in adults and children (2 months of age and older). Several other new therapeutic strategies are also currently under study in addition to the aforementioned molecules (e.g., small-molecule splicing modifiers, anti-myostatin agent, and troponin activator) (7). The current availability of such therapies actually requires the scientific world involved in SMA to rapidly identify reliable measures aimed at quantifying and understanding the response to treatment.

Functional scales such as Revised Upper Limb Module (RULM) (8), Hammersmith Functional Motor Scale Expanded (HFMSE) (9), and 6-Minute Walk Test (6MWT) (10) are commonly used to assess the clinical status and to grade the clinical disability in SMA, classifying patients into four different categories (11, 12). *In vivo*, nerve conduction and in particular compound muscle action potential (C-MAP) have been proposed to quantify the degeneration of αMNs (13). Serum levels of SMN protein have also been shown to correlate both with clinical severity and with C-MAP results. Alves et al. (14) have recently demonstrated that serum creatinine level correlates with both C-MAP results and disease severity, making it a candidate biomarker for SMA progression.

Magnetic resonance imaging (MRI) represents an extremely valuable tool that allows non-invasive assessment of both muscle and motor neurons degeneration. Muscle MRI is increasingly used to evaluate muscle involvement in neuromuscular disorders, to identify specific patterns of disease, to guide biopsy, to monitor the progression of the natural disease course and to assess response to therapy (15, 16). In SMA, muscular involvement has been described as a combination of muscle atrophy and muscle fatty substitution. Durmus et al. (17) reported a specific pattern of muscle involvement in 25 subjects with genetically confirmed SMA 3b, with prominent fat replacement evident in the iliopsoas and gluteus maximus muscles and in the triceps and biceps brachii muscles in the upper limbs. In a recent longitudinal study on two siblings with SMA 3b

Abbreviations: 6MWT, 6-minute walk test; AD, axial diffusivity; aMNs, alpha motor neurons; BL, baseline; C-MAP, compound muscle action potential; CSA, cross-sectional area; DHs, dorsal horns; DTI, diffusion tensor imaging; EPG, extended phase graph; FA, fractional anisotropy; FF, fat fraction; FG, fasciculus gracilis; GM, gray matter; HC, healthy controls; HFMSE, Hammersmith functional motor scale expanded; IZ, intermediate zone; lCST, lateral corticospinal tract; MD, mean diffusivity; MRI, magnetic resonance imaging; MT, magnetization transfer; MTsat, magnetization transfer saturation; qMRI, quantitative MRI; RD, radial diffusivity; ROIs, regions of interest; RULM, revised upper limb module; SC, spinal cord; SMA, spinal muscular atrophy; SMN1, survival motor neuron gene 1; T1w, T1-weighted; T2*w, T2*-weighted; T2w, T2-weighted; TP#, time point #; vCST, ventral corticospinal tract; VHs, ventral horns; w-T2, water T2; WM, white matter.

under treatment with nusinersen, fat replacement was assessed using a semi-quantitative method (18) and no significant changes were demonstrated at 10- and 24-month follow-up, presumably due to the elevated degree of fat replacement already present at baseline (19). Only one recent study explored quantitative muscle MRI measurements in SMA 2 and 3 subjects, showing a significant increase of fat fraction (FF) in SMA subjects compared to healthy controls (HCs) (47.6 vs. 7.6%) by using a four-point Dixon sequence, and investigating also water T2 (w-T2) in muscle (20). Two studies already explored the diffusion properties of muscle tissues in SMA by applying diffusion tensor imaging (DTI) cross-sectionally and longitudinally (19, 20), with promising results. To our knowledge, however, longitudinal quantitative muscle MRI data, in particular FF and w-T2, have never been reported in SMA patients undergoing therapy.

Quantitative MRI (qMRI) of the cervical SC has been widely used to investigate cord pathology in several neurodegenerative disorders such as MS (21, 22) and ALS (23, 24). Through the acquisition of images with high contrast between white matter (WM) and gray matter (GM), SC MRI enables tissue segmentation and measurement of the cross-sectional area (CSA) of the SC, total GM, and ventral horns (VHs), thus making it possible to quantify total SC and GM atrophy (25). Advanced techniques such as magnetization transfer (MT) and diffusion tensor imaging (DTI) of the SC can be used for *in vivo* quantification of possible pathology-related microstructural changes occurring within the SC (26). qMRI of the cervical SC has been recently applied in three recent cross-sectional studies focusing on pediatric and adult SMA patients not under therapy (27–29). El Mendili et al. (27) reported a significant cord atrophy gradient with a more evident reduction of SC CSA between C3 and C6 levels of adult SMA patients compared to HCs; Querin et al. (28) investigated both WM and GM, showing a significant reduction of the CSA of both total SC and GM (but not of WM) at each cervical level (from C2 to C7) in adult untreated SMA patients (SMA 3a, 3b, and 4); finally, in the third study, both pediatric and adult SMA patients (type 2 and 3) showed lower SC CSA values than HCs, though the difference was not statistically significant (29). Two of the studies mentioned above also explored the diffusion properties of SC in SMA: Stam et al. (29) reported significantly higher axial diffusivity (AD) in the GM of SMA patients than HCs, while Querin et al. (28) found preserved diffusion properties of SC WM. However, to date and to the best of our knowledge, there are no imaging studies that longitudinally describe the natural evolution of SC features in SMA patients or the effects of treatment on SC properties measured by qMRI over time.

The aim of this pilot study is to apply muscle and spinal qMRI in a small cohort of SMA 3a patients currently being treated with nusinersen at our Institution, in order to establish whether quantitative muscle and cervical SC MRI parameters can provide biomarkers to monitor the course of the disease, understand possible mechanisms of progression or recovery and assess the effectiveness of treatment.

METHODS

Subjects

In this study we enrolled three female patients affected by genetically diagnosed Type 3a SMA [age at the time of baseline examination—time point 1 (TP1) = 23.0 ± 3.5 years]. Our center provides nusinersen for patients with SMA using intrathecal injections via lumbar puncture, and the dosing schedule consists of four loading doses of 12 mg in the first 2 months of treatment, followed by maintenance doses every 4 months (30). All three patients started treatment at the same time, and the duration of treatment at the time of the last observation time point reported here (TP6) was 21 months. Patients underwent clinical examination at the first loading dose of nusinersen (TP1) and at the time of each maintenance dose. Muscle MRI started to be performed from TP1 as well and was repeated after 9 months (TP3) and 21 months (TP6) from the beginning of therapy. SC MRI examination started at TP3 (9 months after the beginning of treatment, age at TP3 = 24.0 ± 3.5 years) for a total of four time points, each at 4 months from the previous one (TP3 to TP6). Eleven female healthy volunteers, age-matched to that of the SMA patients at TP3 (age 24.5 ± 2.0 years), were also enrolled as HCs for SC MRI only, undergoing a single MRI examination. The study was approved by the local Ethics Committee and written informed consent was given by all participants.

Clinical Evaluation

At each time point, a neurological examination was performed, and the motor function was assessed through the HFMSE and the RULM tests. The 6MWT was also performed in the only ambulatory patient (Patient 1).

MRI Acquisition

MR images were acquired with a Siemens 3T MAGNETOM Skyra scanner using an 18-channel surface coil for the thigh muscle MRI, while for MRI of the SC, a 20-channel head/neck coil was combined with the 12-channel spine array integrated into the patient table.

MRI of the Thigh Muscles

The muscle MRI protocol was centered on the thighs, with simultaneous acquisition of both sides (total scanning time of approximately 15 min). The MRI protocol included a 3D six-point multi-echo gradient echo (GRE) sequence with interleaved echo sampling (matrix size = 432 × 396 × 52, TR = 35 ms, 6 echo times, TE = 1.7–9.2 ms, ΔTE = 1.5 ms, resolution = 1.0 × 1.0 × 5.0 mm^3) and a turbo spin echo T2 multi-echo sequence (TE = 10.9 ms, TR = 4100.0 ms, resolution = 1.2 × 1.2 × 10.0 mm^3, 17 echo times).

MRI of the Cervical SC

The SC MRI protocol adopted for this study is derived from the consensus acquisition protocol produced by a consortium of SC researchers [(31); Cohen-Adad et al., Submitted, https://spine-generic.readthedocs.io/en/latest/]. The whole protocol included T1-weighted (sagittal 3D MPRAGE, field of view = 320 × 260 × 192 mm^3, resolution = 1.0 × 1.0 × 1.0 mm^3, TI/TR/TE

= 1,000/2,000/3.72 ms) and T2-weighted 3D morphological sequences (sagittal 3D SPACE, field of view = 256 × 256 × 52 mm^3, resolution = 0.8 × 0.8 × 0.8 mm^3, TR/TE = 1,500/120 ms), with the T1w volume providing full coverage of the brain and the T2w volume centered on the cervical SC at the level of the C3–C4 vertebral disc. These were followed by a DWI acquisition (axial 2D SE-EPI, field of view = 86 × 33 × 75 mm^3, resolution = 0.9 × 0.9 × 5.0 mm^3, TR/TE = 620/60 ms, b = 800 s/mm^2, 30 directions) using the Siemens ZOOMit technology for reduced FOV (32) and cardiac gating a set of three 3D gradient echo sequences for MT imaging [axial, field of view = 230 × 230 × 110 mm^3, resolution = 0.9 × 0.9 × 5.0 mm^3, TR/TE = 35(15)/3.13 ms], and a 2D T2*-weighted (axial 2D MEDIC, field of view = 224 × 224 × 75 mm^3, resolution = 0.5 × 0.5 × 5.0 mm^3, TR/TE = 600/14 ms) scan, all centered at the level of C3–C4 vertebral disc. The total acquisition time for the cervical SC MRI was approximately 30 min, depending on heart rate.

Analysis of MR Images

MRI of the Thigh Muscle

A total of 12 muscle regions of interest (ROIs) were manually drawn by a single experienced operator using ITK-snap v3.0 (33). These ROIs were drawn at the medial level of the thigh (equidistant from the femur head and the tip of the patella), on two central slices of the first echo of the turbo spin echo sequence. ROIs were subsequently registered to the T2 GRE space. Given the considerable fat replacement of the subjects, a unique global ROI including all muscles and fascia was also drawn independently from the previous ROIs for separate evaluation (Global ROI). **Figure 2** reports muscle segmentation of an exemplificative subject (Patient 2 at TP3). The segmented muscles are reported in the figure caption.

The Fatty Riot algorithm was used offline for the calculation of fat/water images from the GRE multi-echo acquisition (34, 35), and then fat fraction maps were obtained.

The images acquired with the multi-echo spin-echo sequence were processed with an extended phase graph (EPG) method for multi-compartment T2 fitting with slice profile correction: this produced a quantitative w-T2 map of thigh muscles (36–38).

Average values of FF and w-T2 were calculated for the global ROI and for each thigh muscle (see **Figure 2**) and, at a later stage, also for each thigh compartment (anterior, medial, and posterior) (39) at TP1, TP3, and at the end of the follow-up (TP6). Given the very small number of subjects, data were qualitatively compared to assess average FF and w-T2 at the different considered time points.

MRI of the Cervical SC

All MR images of the cervical SC were processed using the Spinal Cord Toolbox (40) and custom scripts in Matlab (The MathWorks, Inc., Natick, MA, USA). Results were visually inspected by a physicist and a trained neuroradiologist, blinded to clinical data, who evaluated the overall image quality and the presence of artifacts; they also manually corrected segmentation masks and vertebral labeling where an unsatisfactory performance or even a failure of the automatic algorithm made it necessary.

T2w images were used to compute the CSA of the SC from C2 to D3. The SC was segmented using a method based on neural networks (41); it was then manually labeled at the different vertebral levels and registered to the PAM50 template (42).

T2*w images were used to compute the CSA of the SC GM and of the VHs from C3 to C4. GM was segmented using a neural network method (43); the template of the SC was registered to T2*w images to obtain vertebral labeling and masks of the VH that were intersected with the GM mask. In order to assess possible asymmetries, GM and VH binary masks were also split into left and right side masks using the corresponding atlas labels that were intersected with the respective masks of interest.

T1w images served as anatomical reference and for diagnostic purposes; they were also used to improve the results of template registration with DWI and MT images.

DWI images were corrected for motion and the SC was segmented (41). They were registered to the SC template and the DTI model was fitted to the diffusion data (44). Warped atlas probabilistic labels were used to extract weighted average values of diffusion tensor metrics (fractional anisotropy, FA; mean diffusivity, MD; radial diffusivity, RD; axial diffusivity, AD) for tissues and structures of interest.

The MT set of images was processed to compute magnetization transfer saturation (MTsat) maps to evaluate the SC myelin content (45). Processing included image co-registration, cord segmentation, and image registration to the cord template. Warped atlas probabilistic labels were used to extract weighted average MTsat values of specific SC regions.

SC CSA values were averaged along the whole considered portion of the SC (from C2 to D3), along cervical and dorsal vertebrae separately, and also per vertebral level. Global DTI and MT average values were obtained from ROIs covering C3 and C4 vertebral levels. The ROIs considered for DTI and MT analysis corresponded to cord GM and its subregions [VHs, dorsal horns (DHs), and GM intermediate zone (IZ)] and cord global WM together with three WM bundles: the ventral corticospinal tract (vCST), the lateral corticospinal tract (lCST), and the fasciculus gracilis (FG). The sample standard deviation was calculated for each average value. The resulting CSA, diffusion, and MT values were compared between HC and SMA groups at the first SC MRI time point (TP3) and also longitudinally between successive time points for SMA patients. Given the small number of SMA patients involved in this study, only qualitative comparisons were performed of each patient quantitative metric against the distribution of values obtained in the HC group.

RESULTS

Clinical Evaluation

Clinical features of the cohort are summarized in **Table 1**. All patients showed the typical clinical picture of SMA (hypotonia, weakness, areflexia, fasciculations, ligamentous laxity in hands) and a slowly progressive clinical course of the disease. Notably, Patient 2 had developed a striking asymmetry in strength of upper limbs since the age of 17. Before starting the treatment, none of the patients needed mechanical ventilation, and only Patient 3 developed a moderate scoliosis not requiring

TABLE 1 | Main clinical features of the enrolled patients.

	Patient 1	Patient 2	Patient 3
Age at BL (years)	27	21	21
Age at onset (months)	24–36	12–24	31
Family history	-	-	One maternal uncle affected by genetically proven 5q SMA
Symptoms at onset	Walking impairment, frequent falls, problems in rising from the floor	Asked to be carried, used her upper limbs to better move lower limbs	Waddling gait
Clinical peculiarities	-	Strength asymmetry in upper limbs (left < right)	-
Age at diagnosis (years)	6 (molecular diagnosis)	3	3
SMN2 copies (*n*)	4	3	4
Age at loss of ambulation (years)	- (still ambulatory)	16	19*
Respiratory problems/scoliosis	No/No	No/No	No/Yes (no surgery)
Treatment duration at TP6 (months)	21	21	21

BL, baseline.

**Patient 3 started using a wheelchair for outdoor at 4; she was still able to move a few steps with support until the age of 11.*

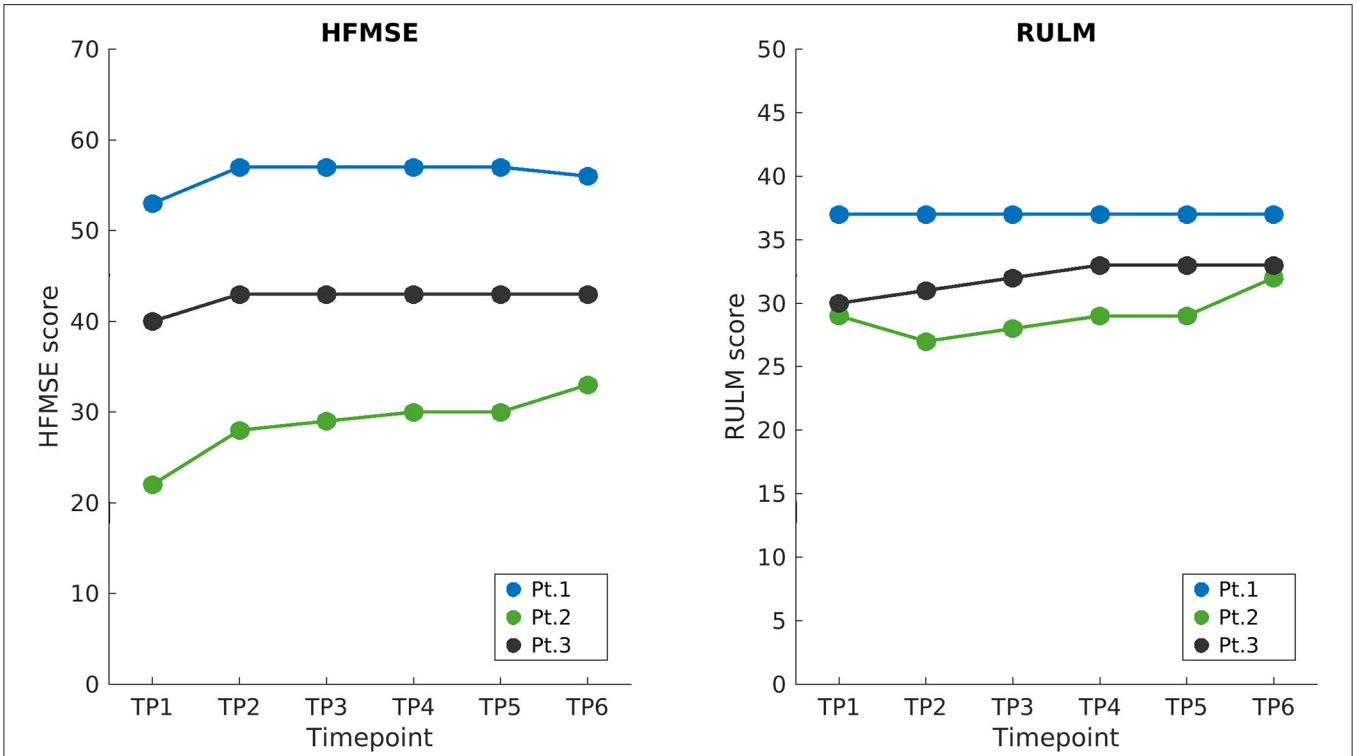


FIGURE 1 | Longitudinal clinical scores. HFMSE and RULM scores of SMA patients are shown on the left and right, respectively, starting from the baseline when treatment began (TP1), up to the end of the follow-up 21 months later (TP6). Loading doses of nusinersen were administered between TP1 and TP2 and the following maintenance doses corresponded to displayed time points (TP2 to TP6) 4 months apart. HFMSE, Hammersmith functional motor scale expanded; RULM, revised upper limb module; BL, baseline; TP, time point.

spine surgery; such features have not changed over the time of observation.

The overview of the longitudinal data of the aforementioned standardized scales is summarized in **Figure 1**. Since the beginning of therapy, all three patients showed a mild improvement in HFMSE scores in the initial period of observation, with a continuous positive trend being only shown for Patient 2. Patients 2 and 3 improved RULM scores at TP6, while Patient 1 remained stable. Additionally, Patient 2, who had a pre-treatment asymmetric strength deficit of the upper limbs (see **Table 1**), showed a bilateral improvement during the observation period (as revealed by the RULM scale). The

TABLE 2 | Per-compartment quantitative muscle MRI measures.

	Anterior compartment	Medial compartment	Posterior compartment	Global ROI
(A) Average Fat fraction (FF) percentage for each compartment (anterior, medial, and posterior) and for the global ROI at baseline (TP1) and at the end of follow-up (TP6). Left and right sides are averaged. The increase/decrease (delta) is reported as percentage.				
TP1	51.2	54.3	45.8	50
TP6	57.1	57.1	54.5	57
Delta	+10.9%	+5.1%	+19.8%	+14%
(B) Average w-T2 in ms for each compartment (anterior, medial, and posterior) and for the global ROI at baseline (TP1) and at the end of the follow-up (TP6). Left and right sides are averaged. The increase/decrease (delta) is reported as percentage.				
TP1	43.7	44.4	30.4	43.73
TP6	42.6	41.1	28.1	41.68
Delta	−2.4%	−7.2%	−0.6%	−4.7%

performance of Patient 1 in 6MWT remained constant over time points.

Muscle MRI

The image quality was reviewed by an experienced radiologist and positively evaluated for subsequent analysis.

All data for individual muscles and global ROIs are reported for TP1, TP3, and TP6 in the **Supplementary Tables 1A,B**, while **Table 2** show the FF and w-T2 and their evolution over time for each compartment separately and globally (global ROI) at TP1 and at the end of the follow-up (TP6).

The average value of FF across the three SMA patients evaluated with a single ROI (see **Figure 2**) was 50% at TP1 and increased to 56.7% at the end of the follow-up (TP6). All muscles showed progression of FF with the only exception of the adductor longus (AL), which showed a reduction of FF from 39 to 37% (−6.02%) at TP6 (**Supplementary Table 1A**). For this reason, we decided to exclude AL from the analysis of the medial compartment and considered it on its own. In the posterior compartment, the most involved muscle was the semitendinosus (ST) (0.47% FF at TP1), which also progressed to 0.77% at TP6.

For the anterior compartment, FF progressed from 51.2 to 57.1% (average increase of +10.87%); for the medial compartment (excluding AL), from 54.3 to 57.1% (+5.12%); and for the posterior compartment, from 45.8 to 54.5% (+19.81%).

For what concerns the evaluation of w-T2, we found that w-T2 on the global ROI was 43.73 ms at TP1 (the higher value being represented by the medial compartment), decreasing over time (−4.7% in average at TP6 compared to TP1). All muscles showed a decrease of w-T2 over time (from TP1 to TP6), with the exception of AL, which showed an increase over time of 6% (from 43.84 to 45.11 ms) (**Supplementary Table 1B**).

SC MRI

Dataset and Image Quality

Images from an exemplary dataset of a randomly selected HC are shown in **Figure 3**.

The image quality was overall very good. Where motion artifacts and loss of signal and contrast in 3D T1w and T2w images were detected, we interrupted segmentation and vertebral labeling below their D1 and D2 levels, respectively, for one HC, and below C6 for Patient 3 at TP5.

Cross-Sectional Analysis of SC

Figure 4 reports group results of CSA values evaluated on T2w images of HC and SMA subjects acquired at TP3. We found that SC CSA mean values of SMA patients fall in the lower range of the HC distribution, with two patients lying below the 25th percentile. The "per level" analysis additionally showed that the difference between the two groups was more pronounced at the cervical level with respect to upper dorsal vertebrae, even if the cervical intra-group variability is higher than the dorsal one for both groups; the difference between HC and SMA is maximum around from C3 to C6. SC CSA values computed on T2*w and T1w images confirmed this result.

CSA of GM and VHs analyzed on T2*w images were reduced in patients compared to HC. No noticeable asymmetry was found between the right and left sides of GM and VHs (see **Figure 5**).

Figures 6, **7** report group results of DTI measures (FA and RD, respectively) for HC and SMA subjects at TP3. Analogous plots for AD and MD can be found in **Supplementary Figures 1**, **2**. In GM, we report a reduction in RD and an increase in FA in SMA patients with respect to HC, while no differences can be observed for AD and MD values. Accordingly, FA and RD values of the VHs of two out of three SMA patients fall outside the range of the corresponding HC distribution, while MD and AD of the VHs differ from the range of HC to a lesser extent.

We found no univocal variation in WM and that DTI values of most WM regions in SMA patients fell within the range of the respective HC distribution.

Figure 8 reports group results obtained from MTsat maps of HC and SMA patients at TP3. We found a decrease of MTsat values in both WM and GM of SMA patients with respect to the HC group. Moreover, we found an evident decrease of MTsat average values in the specific investigated WM bundles (vCST, lCST, and FG). In particular, we found that the MTsat values of two patients out of three fell always outside the corresponding HC group range of values for each of these regions. In GM, the reduction of MTsat is more pronounced in the intermediate regions of the GM, with respect to ventral or dorsal horns.

Longitudinal Analysis of SC

We report a trend of increasing SC, GM, and VHs CSA values in T2*w images of SMA Patient 2. CSA values of Patients 1 and 3 do not show evident trends over time (see **Figure 9**). Mean DTI metrics computed do not show evident trends over time and they are consistent between patients (see **Supplementary Figures 3–6**). The longitudinal trend observed for MTsat values shows great similarities between different ROIs. In particular, this longitudinal analysis highlighted a substantially stable pattern of MTsat in each ROI of Patient 2. On the other hand, data from Patient 1 showed a remarkable increase from TP3 to TP4, while Patient 3 exhibited an abrupt drop from TP5 to TP6 (see **Figure 10**).

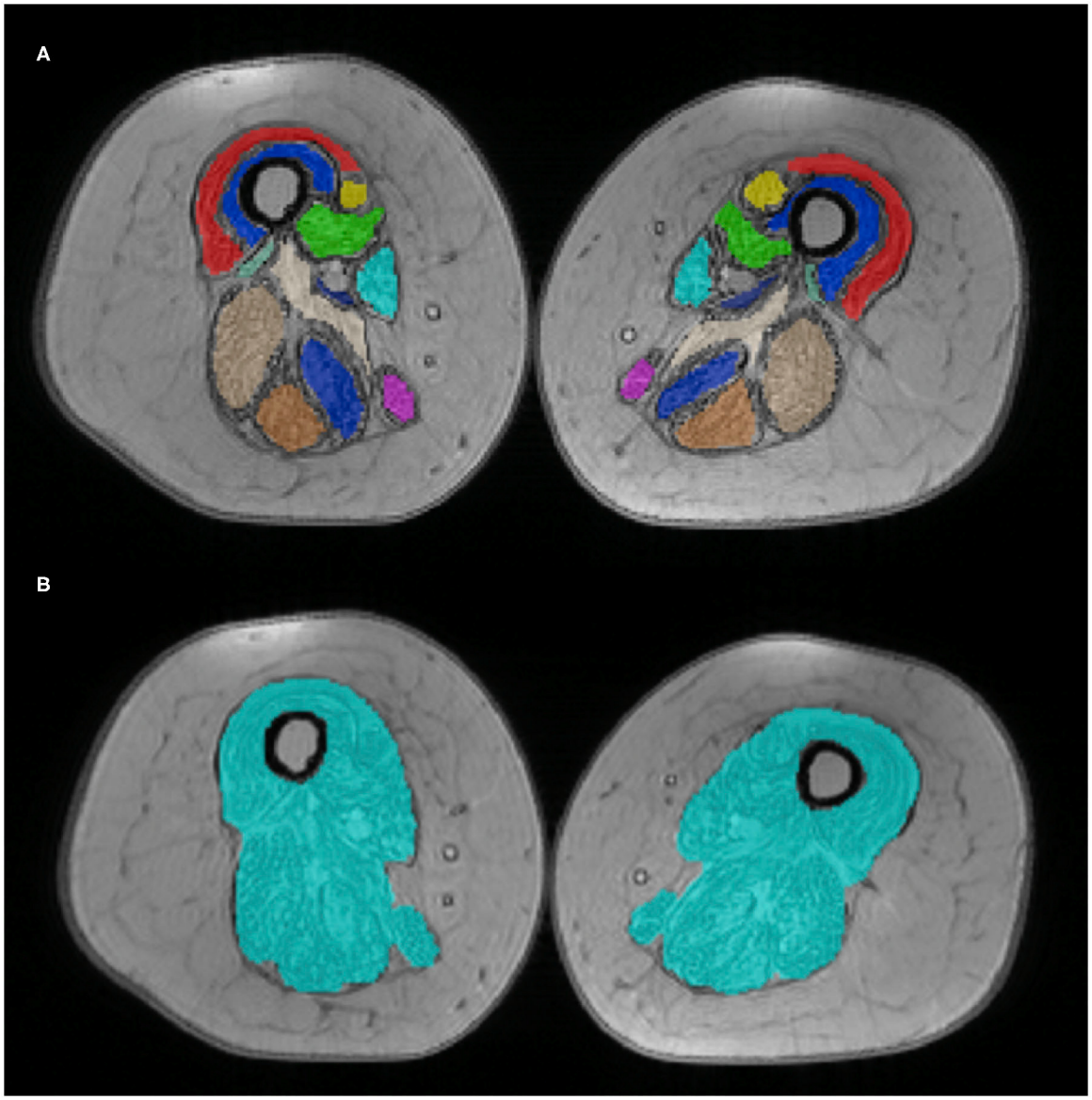


FIGURE 2 | Thigh muscle regions of interest. Example of manually drawn muscle ROIs at middle thigh level, superimposed on the T2 GRE sequence (Patient 2 at TP1). On top **(A)**, the 12 muscle ROIs are individually displayed with different colors; at the bottom **(B)**, the global ROI comprising all muscles is shown. The 12 ROIs include VL, vastus lateralis; VM, vastus medialis; VI, vastus intermedius; RF, rectus femoralis; S, Sartorius; G, gracilis; AM, adductor magnus; AL, adductor longus; SM, semimembranous; SM, semitendinosus; BFL, biceps femoris longus head.

DISCUSSION

This pilot study applied qMRI to measure changes in muscle and cervical SC of three adult Type 3a SMA patients under nusinersen treatment with a longitudinal follow-up of 21 months, assessing muscle and SC degeneration, with the aim to evaluate whether MRI could provide useful biomarkers to monitor the clinical efficacy of nusinersen. Patients were clinically stable over the course of this study, while muscle qMRI suggested the presence of fat substitution progression and w-T2 slight decrease over time in thigh muscles. In the SC, we observed reduced CSA, GM, and VHs CSA of SMA patients transversally compared to HCs at TP3; these values remained generally stable over time, although Patient 2 exhibited a trend of increasing GM and VHs area over time. Diffusion parameters were altered when compared to HCs, but stable over time in SMA subjects, while MTsat values of SC structures were also reduced.

Clinical Data

Clinically, all enrolled SMA Type 3a patients were young adult females with similar features, characterized by the early onset of a typical clinical picture of SMA, and a slow progression.

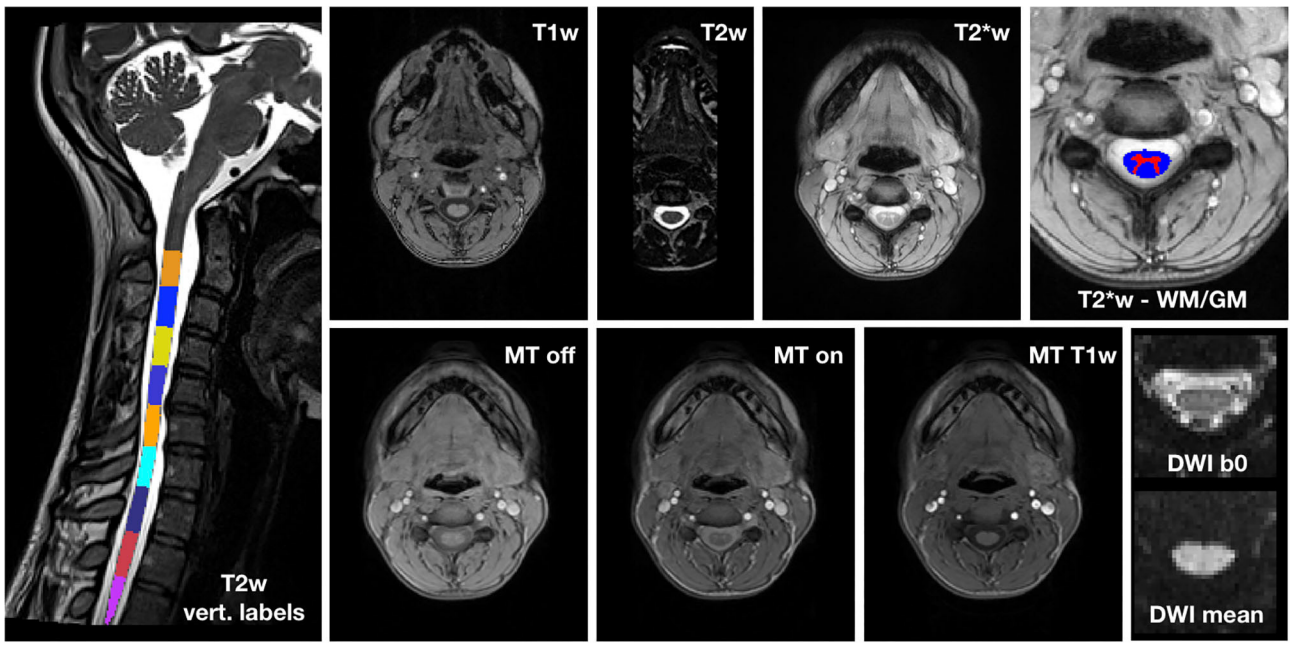


FIGURE 3 | Exemplary spinal cord MRI dataset. On the left, T2w image with vertebral labels from C2 to D3. Top row: T1w image; T2w image; T2*w image; detail of T2*w image with WM and GM segmentation masks. Bottom row: MT (MT-off, MT-on and MT-T1w) and DWI (mean b0 image and mean DWI) dataset.

The three patients reported a subjective general improvement as to their motor performance and "self-confidence" during the study; none of them experienced any side effect during treatment. This subjective perception of improvement in motor performance reported by patients was also confirmed by the results of validated motor function tests: all patients obtained an improvement in the HMFSE scale from TP1 to TP2, when they reached a plateau. Patients 2 and 3 exhibited also an improvement in the RULM score at TP6, while the score of Patient 1 remained stable. This could be due to a general better motor function of Patient 1, who had a very high score already at the baseline (TP1). Furthermore, Patient 2, who had developed asymmetric weakness in upper limb scores before the treatment, showed a bilateral good recovery, but asymmetry essentially remained identical. The normal cervical MRI (no signs of compression or plexopathy were found) and the absence of both a clinical history of trauma and other neurological signs suggest that such asymmetry could be considered as an atypical manifestation of the disease in this patient. Even though SMA is generally described as a symmetric disorder, it is not uncommon to see asymmetrical weakness: Kang et al. (46) reported three female type 3 SMA patients with asymmetrical weakness, showing also upper motor neuron signs that have always been absent in our patient. Despite the small sample size considered here and the relatively short observation period, our results are in line with those reported by the German and Italian studies (47, 48). It should also be noted that all patients reported some further improvement in motor function, which could not be captured by their repeated functional scores possibly because of the discrete nature of such scales.

Muscle MRI

In all three patients, a progression of thigh muscle fat substitution was documented over time (mean increase of roughly 7% in 21 months when considering the global cross-sectional muscle ROI). This progressive fat substitution can represent the natural evolution of the disease over time, despite the concomitant therapy and the global stability of clinical scales. Unfortunately, to our knowledge, no longitudinal quantitative muscle MRI data on the natural history of the disease are available for comparison at present. The posterior compartment of the thigh showed the greatest percentage increase in FF (+19.81%) in our cohort, but the most elevated fat replacement was shown in the anterior and medial compartments (both with 57.1% FF at the end of follow-up, TP6), the medial compartment being the most involved from the beginning of the study. This anterior compartment predominant involvement is in line with other qualitative and quantitative muscle MRI studies conducted so far, including the relative sparing of the adductor longus (AL) (26, 39). We also confirmed the greater fat replacement of the semitendinosus (ST) among other hamstring muscles, already demonstrated semiquantitatively by Brogna et al. (39). By contrast, the relative sparing of the medial compartment previously demonstrated semiquantitatively by Brogna and colleagues was not confirmed by the present quantitative data, possibly because we did not perform a muscle CSA analysis.

Given the small number of cases included in this pilot study, we did not perform any correlation analysis between quantitative muscle features and clinical scales, neither at TP1 nor at the end of follow-up. Even considering the global stability of the disease as observed with clinical scores, FF quantification

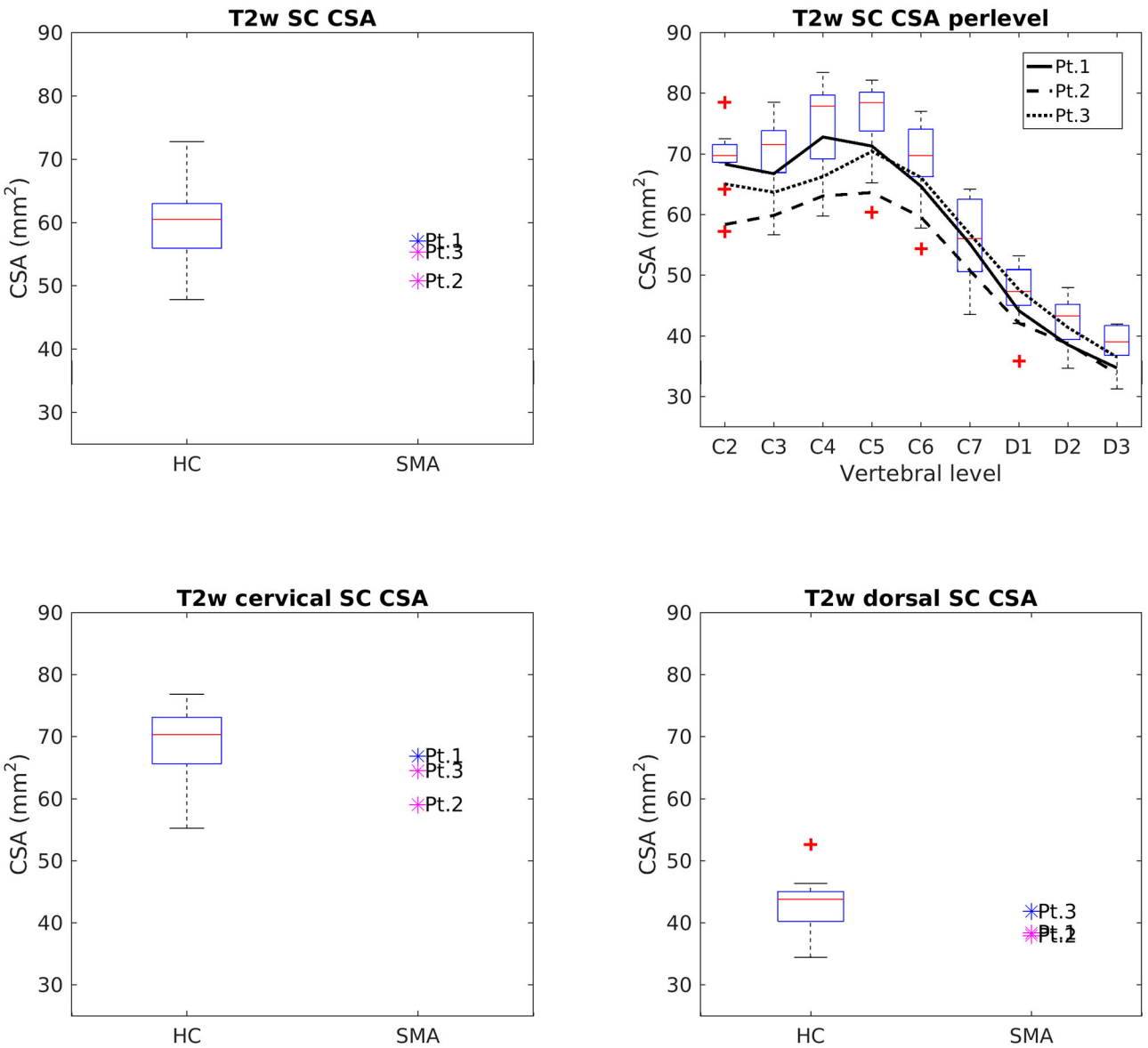


FIGURE 4 | Cross-sectional comparison of CSA values measured on T2w images at TP3. Top left: average SC CSA from C2 to D3 in HC (box on the left) and SMA patients (dots on the right). Top right: per-level SC CSA values from C2 to D3 in HC (boxes) and SMA patients (black lines). Bottom left: average SC CSA from C2 to C7. Bottom right: average SC CSA from D1 to D3. The color of patient dots relates to their position compared to the respective distribution of HC values: blue for patient values within the 25–75th percentile range, magenta for patient values outside the 25–75th percentile range but within the HC distribution, red for patient values outside the HC distribution. Red + represent outlier HCs.

was able to detect subtle changes in the muscle, thus once again demonstrating the potential of applying fat replacement quantification as a surrogate outcome measure for clinical trials.

FF, however, represents only part of the disease process, becoming evident at advanced stages pathology. By contrast, high w-T2 signal has been proposed in neuromuscular disorder studies as a potential marker of disease activity, though non-specific, mostly reflecting initial stages of the disease (15).

Previous studies exploring w-T2 in thigh muscles in SMA patients reported an increased value in thigh muscles (49, 50) compared to HCs. Bonati et al. (49) obtained w-T2 values exceeding 60 ms in SMA muscles, but used a mono-exponential fit for the T2w multi-echo sequence. Chabanon et al. (50) applied a multi-exponential signal model and found increased w-T2 values (ranging from 34.3 to 31.3 ms), but did not take into account EPG fitting. By contrast, Otto et al. (20) adopted a multi-exponential model accounting for EPG fitting and did not report

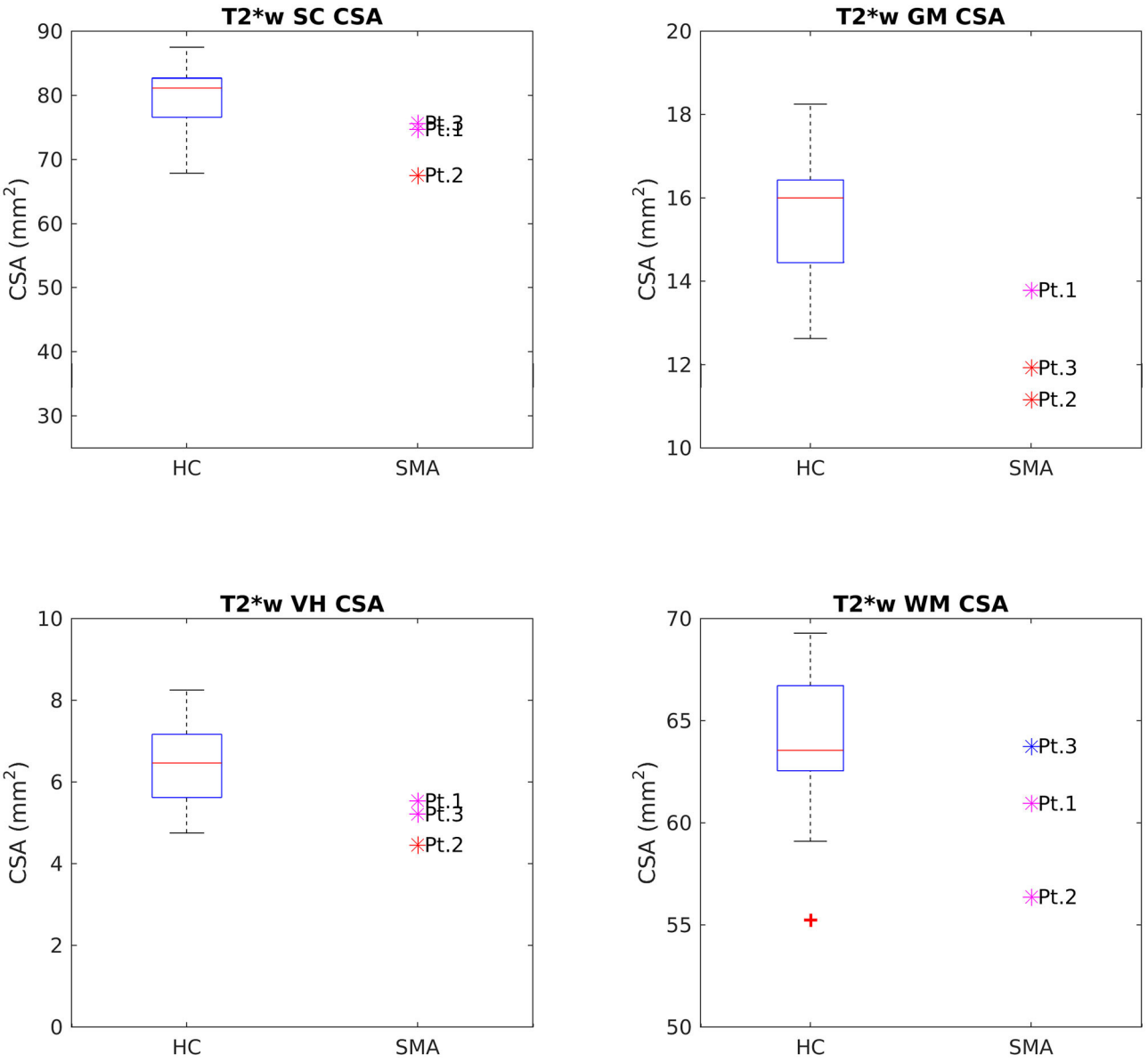


FIGURE 5 | Cross-sectional comparison of CSA values measured on T2*w images at TP3. From left to right and top to bottom: average CSA values of SC, GM, VHs, and WM measured from C3 to C4 in HC (box on the left) and SMA patients (dots on the right). The color coding of patient dots is explained in the caption of **Figure 4**. Red + represent outlier HCs.

values higher than 31 ms at most in SMA, thus suggesting that differences in values could be mostly due to methodological issues and not to pathological mechanisms.

In the present longitudinal study, we also applied multicompartment EPG fitting for T2 mapping and we found thigh muscle w-T2 ranging from 39.66 to 52.64 ms in SMA patients (average 43.73 ±3.73 ms at TP1, before the start of therapy), values that are slightly high when compared to previously reported values (20, 49). Additionally, w-T2 seemed to slightly decrease over time, with the greatest percentage reduction in the medial compartment at the last time point (TP6, 21 months after the beginning of therapy). As already reported by Otto et al. (20), such data have to be considered cautiously and we agree with the authors that artifacts or methodological issues may have a great impact in evaluating w-T2 in highly fat-replaced muscles. In addition to this, we know from the literature that w-T2 values are quite variable and linked to a number of physiological and paraphysiological conditions (in addition to the pathological ones) (15, 51). However, we still cannot say whether such abnormalities are due to the natural evolution of muscle pathology in SMA or therapy-induced. Longitudinal future studies applying quantitative

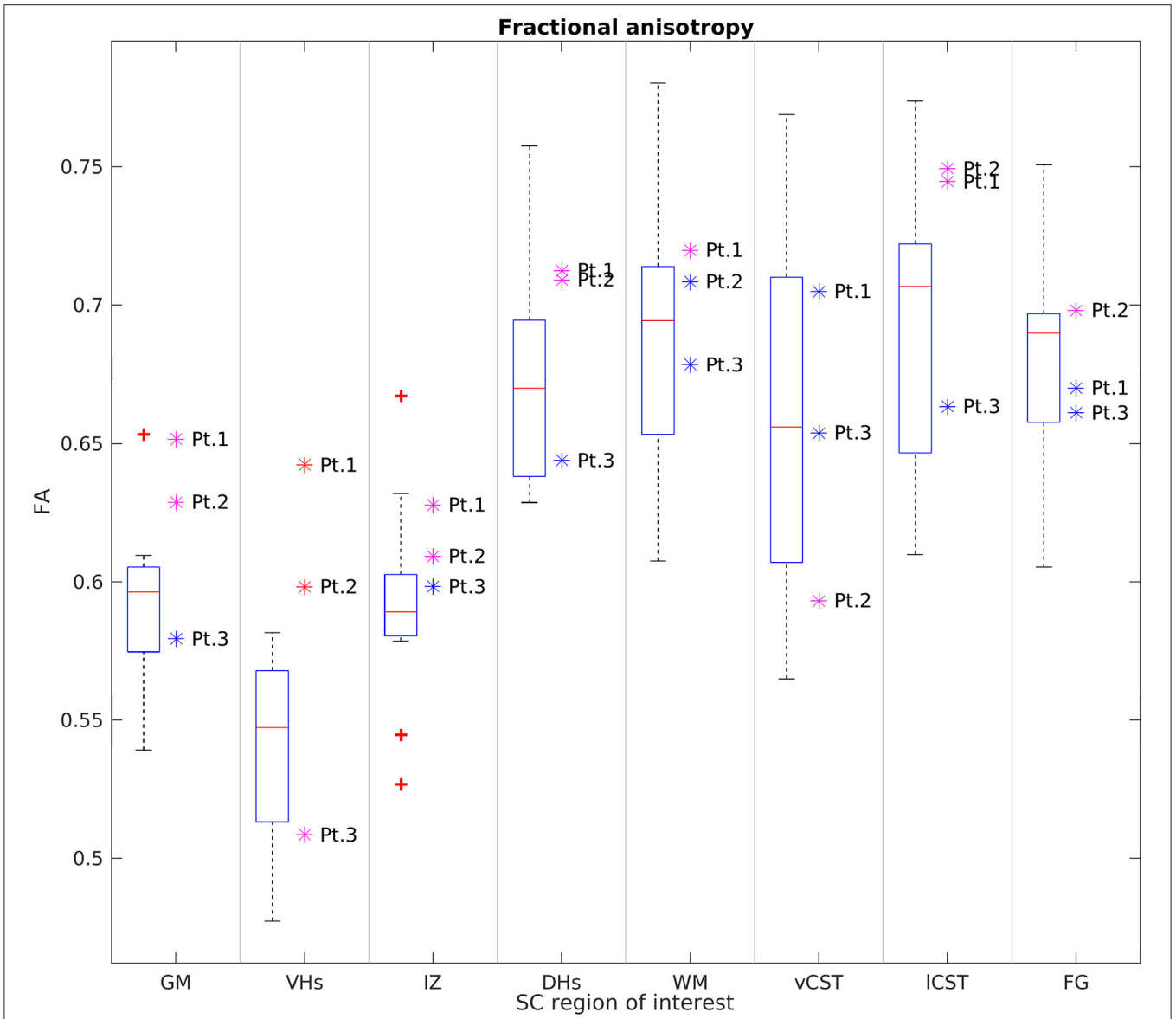


FIGURE 6 | Cross-sectional comparison of average FA values measured at TP3. From left to right: average FA values of GM, VHs, IZ, DHs, WM, vCST, lCST, and FG measured from C3 to C4 in HC (box on the left) and SMA patients (dots on the right). The color coding of patient dots is explained in the caption of **Figure 4**. Red + represent outlier HCs.

muscle MRI to evaluate w-T2 with optimized methodological post-processing and performed on larger cohorts may help to address the question.

DTI has recently been applied as a supplementary advanced muscle MRI technique to evaluate microstructural changes in muscle in SMA: Barp et al. (19) found a reduced fractional anisotropy (FA) over time during nusinersen treatment (24 months follow-up) in two patients, whereas Otto et al. (20) found an increase in FA (and a decrease in MD) in SMA compared to HCs. DTI studies of thigh muscle may help to investigate muscle microstructural properties in SMA beyond FF and w-T2, as already suggested (20).

Cervical SC MRI

Cross-Sectional Analysis of SC

We found a clear difference in total CSA of the SC between SMA patients and HCs at TP3, in substantial agreement with previous studies (27–29). SMA patients showed a smaller CSA than HCs, with the largest discrepancy being observed in correspondence of the cervical vertebrae. Such finding is consistent with the typical limb girdle weakness of type 3 SMA. This result was confirmed by segmentations obtained from both T1w and T2w volumetric sequences; the difference in the absolute values of CSA from the two methods can be explained by the different spatial resolution of the two acquisitions. SC GM was segmented in high in-plane resolution T2*w sequences between C2 and

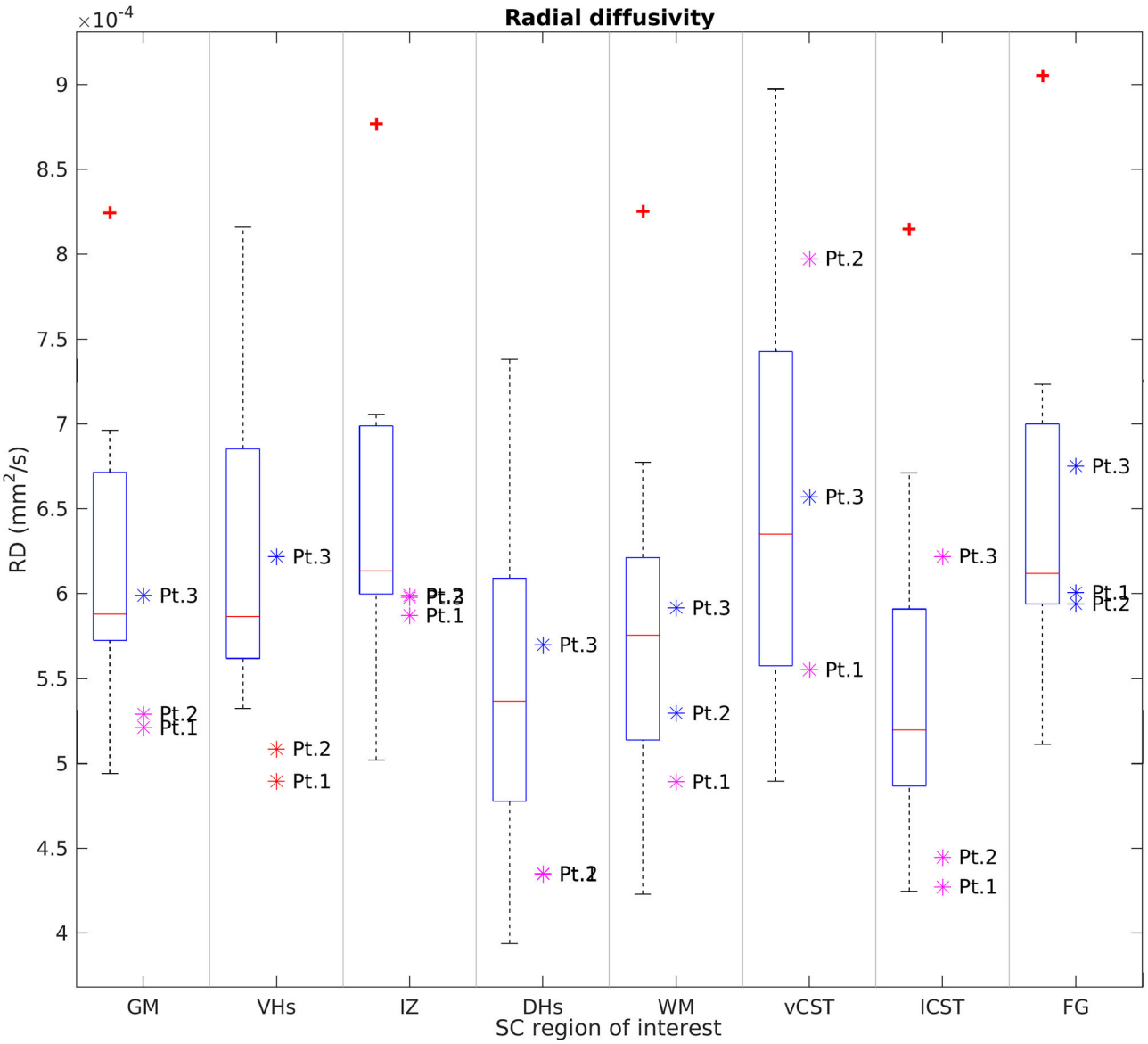


FIGURE 7 | Cross-sectional comparison of average RD values measured at TP3. From left to right: average RD values of GM, VHs, IZ, DHs, WM, vCST, lCST, and FG measured from C3 to C4 in HC (box on the left) and SMA patients (dots on the right). The color coding of patient dots is explained in the caption of **Figure 4**. Red + represent outlier HCs.

C4 and SMA patients showed a reduced area with respect to HCs, thus confirming a pathology-related GM atrophy. The VHs contain the motor neurons population of the SC and, therefore, their shrinkage could reflect specific αMNs degeneration (12). However, given the small size of these structures in relation to the MRI resolution, this result could be affected by partial volume and segmentation errors and should therefore be interpreted with caution.

Even though SMA is generally described as a symmetric disorder, in clinical practice, it is not uncommon to see asymmetrical weakness. In this case, the analysis that we performed to investigate the asymmetry of clinical worsening observed in Patient 2 before beginning treatment produced no evidence of altered qMRI parameters between the two sides that could explain it.

Our measurements of the CSA of SC, GM, and VH are in line with the known degeneration of αMN caused by SMA. In agreement with previous studies (28, 29), we did not detect pathological alterations of diffusion parameters in the WM of SMA patients. Conversely, we documented an increase of GM FA and a slight reduction of RD, while AD and MD were unchanged. In particular, a similar FA variation was also found in VH ROIs. As pointed out also by Stam et al. (29), resolution-driven partial volume effects due to the inclusion

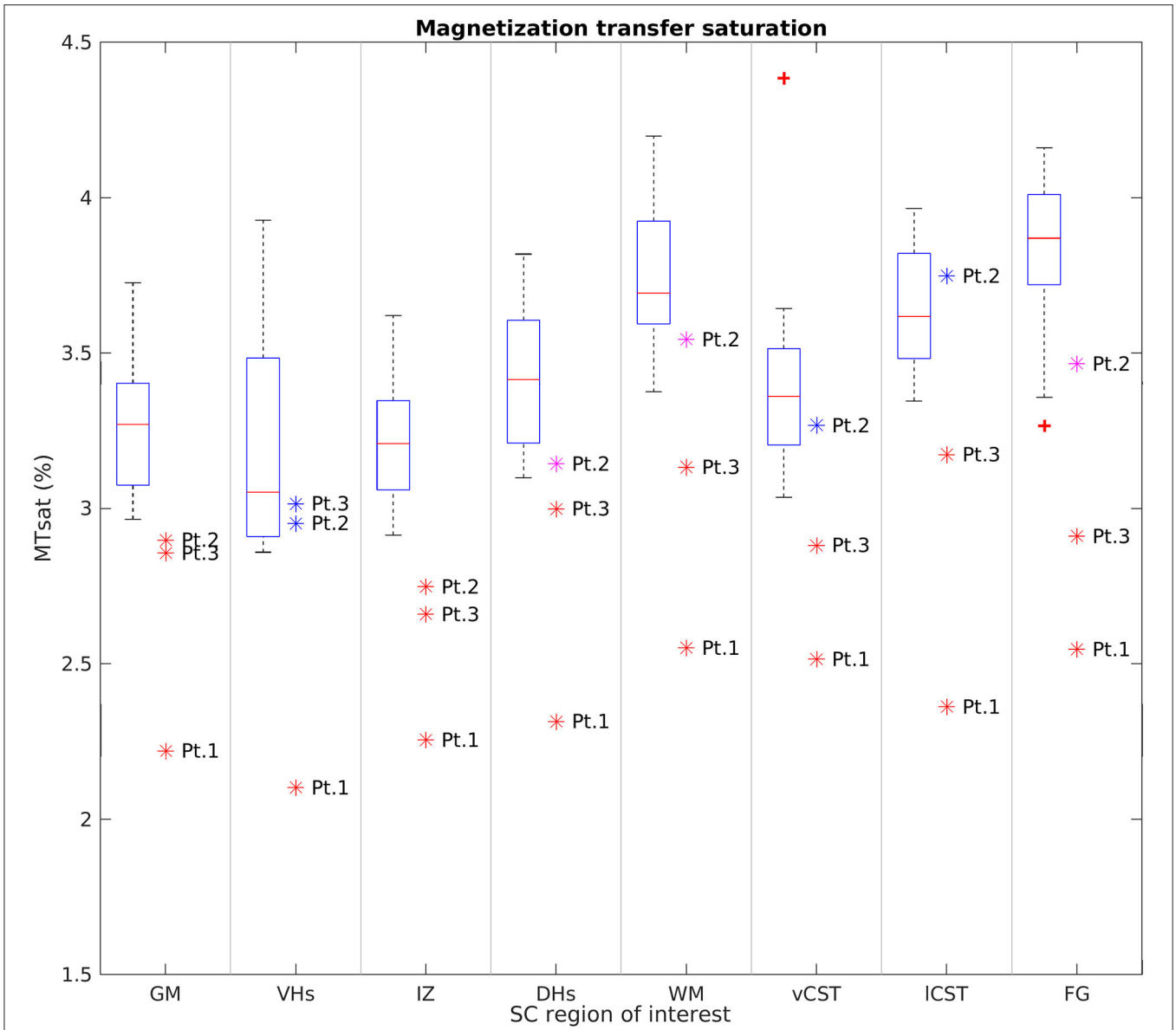


FIGURE 8 | Cross-sectional comparison of average MTsat values measured at TP3. From left to right: average MTsat values of GM, VHs, IZ, DHs, WM, vCST, lCST, and FG measured from C3 to C4 in HC (box on the left) and SMA patients (dots on the right). The color coding of patient dots is explained in the caption of **Figure 4**. Red + represent outlier HCs.

of adjacent WM regions could be a possible cause of altered average FA values, because such effects can be further enhanced by GM shrinkage induced by pathology. However, partial volume effects should be limited by the use of probabilistic labels in the averaging operation. Furthermore, if such changes were due to only partial volume, we should also expect altered MD and AD values because DTI maps are inherently co-registered. Therefore, another possible explanation is that motor neuron degeneration, together with the consequent GM shrinkage, enhances the relative volume fraction of those WM fibers running through the SC GM, thus increasing the resulting FA. A third possibility is that we are observing the effects of gliosis, which was reported to occur in the VH of SMA patients (52). Given the low number of literature histopathological studies, these hypotheses should only be regarded as speculative. Future *in vivo* and *ex vivo* diffusion studies, conducted with newly developed high resolution DWI techniques and fitted with advanced diffusion models, together with concomitant histological studies, will certainly help to better discriminate between these scenarios.

To further understand the microstructural modifications underlying SMA, we acquired MT-weighted images. Changes in MTsat values are considered an indirect marker of changes in myelin content (45, 53). We found a remarkable reduction

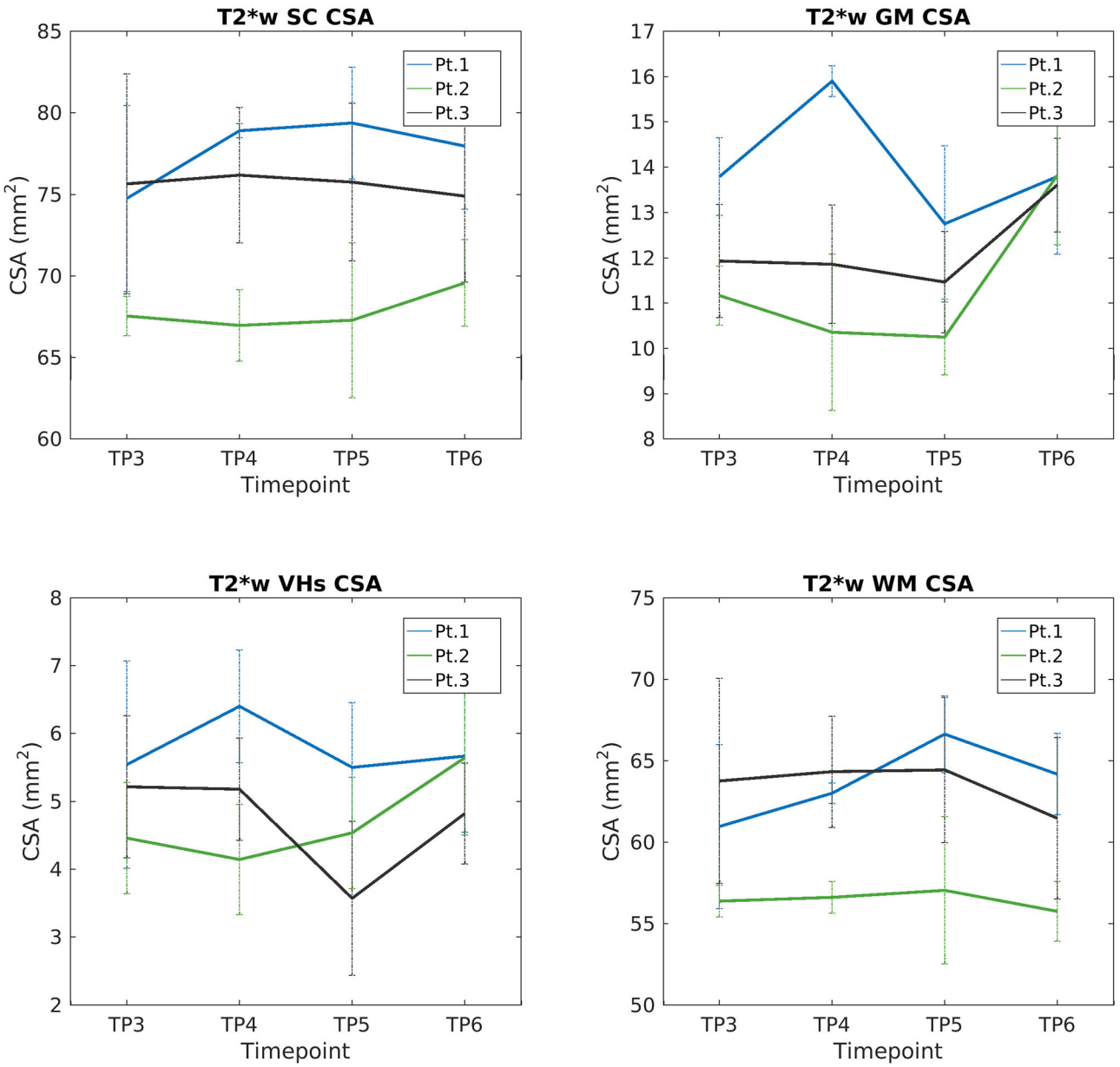


FIGURE 9 | Longitudinal analysis of average CSA values measured on T2*w images. From left to right and top to bottom: average CSA values of SC, GM, VHs, and WM measured from C3 to C4 starting from TP3 to TP6.

of MTsat values in both WM and GM, suggesting a loss of myelinated fibers. The reduced MTsat values observed in FG are in agreement with previous neuropathological studies of SMA, which consistently reported a loss of myelinated fibers in such tracts (52, 54, 55). In addition, myelin loss was also occasionally found to occur in the lateral columns (55, 56) and MTsat values from the corticospinal tracts of SMA patients confirm this. The alteration of MTsat values in SC GM could be explained by partial volume between GM regions and the abovementioned WM tracts; this hypothesis is supported by the fact that the most pronounced reduction of MTsat values was found in the intermediate region of the GM, which is the closest to the FG, the tract that was most frequently reported to be severely affected by demyelination. To the best of our knowledge, this is the first time that MT imaging is applied to investigate SC myelin content in SMA *in vivo*. Given the encouraging agreement between our results and those reported in neuropathological studies, MT measures certainly represent promising candidates for better understanding and monitoring the evolution of SMA.

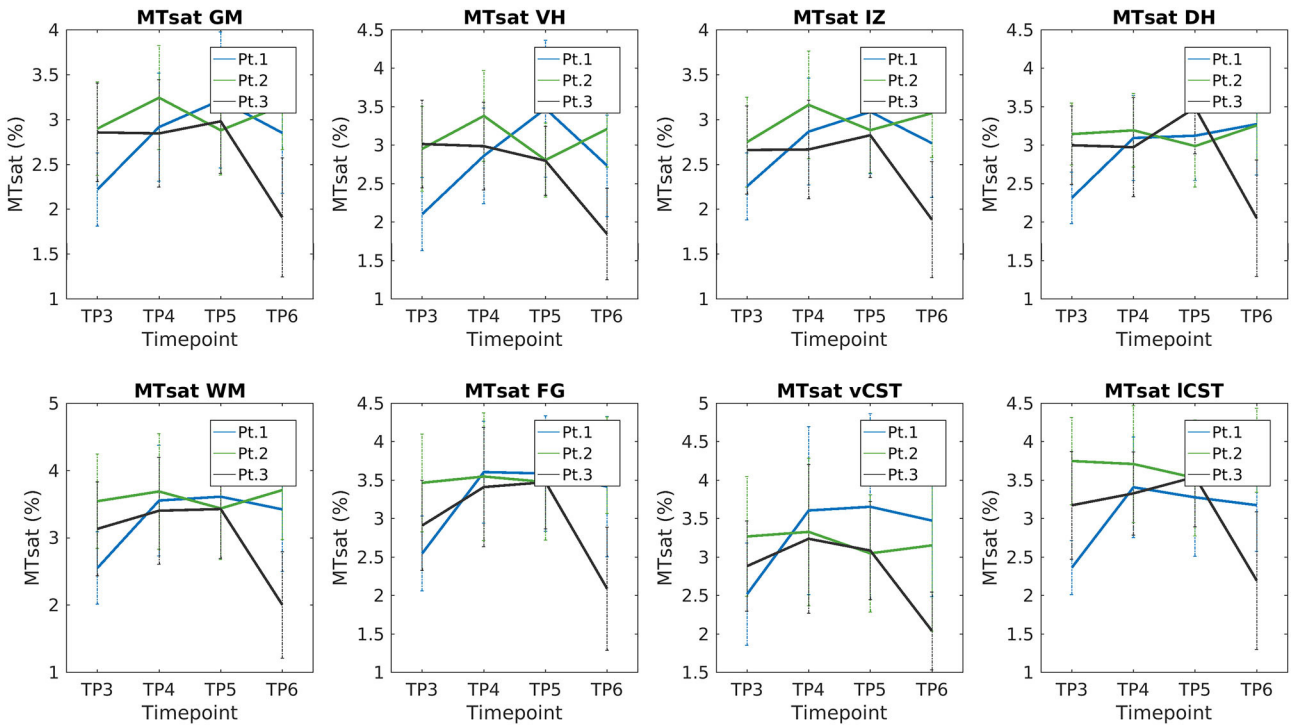


FIGURE 10 | Longitudinal analysis of average MTsat values. From left to right and top to bottom: average MTsat values of GM, VHs, IZ, DHs, WM, vCST, lCST, and FG measured from C3 to C4 starting from TP3 to TP6.

Longitudinal Analysis of SC

The longitudinal evaluation of CSA and GM over time showed stable values for two out of three patients, while Patient 2 showed a slight increase in SC, GM, and VH CSA values. Therefore, the SC volume can be considered stable over 1 year of disease and treatment monitoring. We can hypothesize that the relative volume stability of SC could represent a positive response to therapy: we do not have the pre-treatment quantitative volumetric data of the natural history of the disease and we do not have a control group of untreated SMA patients, yet we could assume that SC volumes (total, GM, and VHs) would have presumably shown a progressive reduction over time without treatment. Such hypothesis is in line with the stability of SC and GM CSA values over time that mirror the clinical improvement shown by increased HMFSE and RULM scores. From the histological point of view, an increased volume of the VH could be compatible with the effects of nusinersen on the motor neuron population, as suggested in a preclinical study in mice (57). The authors used choline acetyltransferase (ChAT, a marker of αMNs) to stain the SC αMNs in a murine model of SMA and reported that nusinersen could effectively increase the number of ChAT-positive cells by 38%, and up to 62%, with respect to untreated SMA mice, thus proving the recovery of a fraction of the degenerated cell population. This result was also confirmed by another group that "detected a recovery in the number and chemical composition of Cajal Bodies" in treated SMA patients, suggesting that the depletion process of Cajal Bodies found in motor neurons from the VH of untreated SMA mice models may be reversed by nusinersen (58). The number of Cajal Bodies is strongly bound to the cell body size (59); therefore, the high number of motor neurons involved in the degeneration/recovery processes and the histological findings reported above suggest that an enlargement of GM, and in particular of the VH areas, could be expected during nusinersen treatment; our results suggest that this effect could also be detected with advanced MRI techniques. It can be argued also that such changes in cell body size during the observation period could impact GM diffusion parameters and, in particular, MD. However, it should be considered that no changes in GM MD were detected between HC and SMA groups at TP3 and that the MD of SMA patients remained stable over time. Consequently, we could hypothesize that the degree of cellular swelling may not be sufficient to produce detectable effects on MD, but it is not excluded that higher-order diffusion models may be able to capture such subtle changes. Furthermore, we do not know exactly the time scale over which these changes may occur after beginning treatment. In this regard, it is essential that future studies closely monitor SC parameters before and during the administration of the loading dose of nusinersen.

Limitations

In this study, we acknowledge a number of limitations. The first and most relevant is the small sample size of the cohort of enrolled patients that impacted our ability to perform statistical analysis. In addition to that, the study is penalized by the lack

of a group of untreated SMA patients: this would have provided a benchmark to better quantify the effects of nusinersen on SC structures over the observation period. Second, we started muscle MRI at TP1, but SC at TP3, so we do not have a proper baseline for this technique. Moreover, it is also possible that 1 year is too short to detect possibly subtle, though significant, changes induced by therapy. To this end, we plan to recall the patients 1 year from the last MRI session to repeat the qMRI scan. These limitations prevented us from clearly assessing whether qMRI parameters can be considered valid biomarkers of disease evolution and treatment efficacy. Future studies should aim to enroll larger cohorts of patients for a longer observation period in order to further investigate the relationship between qMRI parameters and clinical scores through adequate statistical analysis and disease progression models. This could provide a valuable tool to better characterize the natural history of the disease and to predict its evolution in particular in relation to different therapeutic approaches.

A final consideration is that more advanced diffusion and myelin imaging techniques could provide more specific or sensitive biomarkers of pathology. However, we believe that the SC MRI protocol adopted here represents a good trade-off between clinical and research needs, thanks to its relatively short acquisition time in relation to the high number and variety of techniques involved, which allowed us to investigate several aspects of the SC (micro-)structure. On the basis of the findings reported here, it will be possible to make informed choices to better address more specific questions and challenges posed by the disease.

CONCLUSIONS

In this pilot MRI study, we demonstrated a progression of fat substitution in thigh muscles of SMA 3a patients during therapy and a concomitant trend toward a slight reduction of w-T2. The analysis of SC data confirmed a degeneration of tissues highlighted by both GM atrophy and, interestingly, reduced MTsat values, the latter of which have never been reported in SMA before. Our data warrant future studies to further investigate whether qMRI parameters may represent valid biomarkers of pathology and treatment efficacy in SMA, but this will require longer longitudinal studies on larger cohorts and proper statistical models of disease progression.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because of institutional policies. Requests to access the datasets should be directed to anna.pichiecchio@mondino.it.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Comitato Etico Area Referente Pavia della Fondazione IRCCS Policlinico San Matteo. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AP, CG, GS, and MP contributed to the conception and design of the work. GS, FSa, and XD set up the acquisition protocol. GS, FSo, MP, AP, and GG acquired the data. GS, CA, FSo, NB, and SM performed data analysis. GS, AB, CG, SP, AP, MP, and EP contributed to the interpretation of results. AB, AG, SP, and EP performed clinical evaluation of patients. LF, AP, and SB administered the treatment. GS, CA, MP, and AP wrote the first draft of the manuscript. All authors contributed to manuscript critical revision, read, and approved the submitted manuscript.

FUNDING

This work was supported by grants of the Italian Ministry of Health to RCR-2019-23669119_001-Rete RIN Neuroimaging, RC 2017-19, and RC 2020.

ACKNOWLEDGMENTS

We acknowledge the patients with their families for the collaboration in the study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.613834/full#supplementary-material

REFERENCES

1. Van Den Berg-Vos RM, Van Den Berg LH, Visser J, De Visser M, Franssen H, Wokke JHJ. The spectrum of lower motor neuron syndromes. *J. Neurol.* (2003) 250:1279–92. doi: 10.1007/s00415-003-0235-9
2. D'Amico A, Mercuri E, Tiziano FD, Bertini E. Spinal muscular atrophy. *Orphanet J. Rare Dis.* (2011) 6:71. doi: 10.1186/1750-1172-6-71
3. Mercuri E, Bertini E, Iannaccone ST. Childhood spinal muscular atrophy: controversies and challenges. *Lancet Neurol.* (2012) 11:443–52. doi: 10.1016/S1474-4422(12)70061-3
4. Kolb SJ, Kissel JT. Spinal muscular atrophy. *Neurol. Clin.* (2015) 33:831–46. doi: 10.1016/j.ncl.2015.07.004
5. Wadman RI, Stam M, Gijzen M, Lemmink HH, Snoeck IN, Wijngaarde CA, et al. Association of motor milestones, SMN2 copy and outcome in spinal muscular atrophy types 0-4. *J. Neurol. Neurosurg. Psychiatry.* (2017) 88:364–7. doi: 10.1136/jnnp-2016-314292
6. Groen EJN, Talbot K, Gillingwater TH. Advances in therapy for spinal muscular atrophy: promises and challenges. *Nat. Rev. Neurol.* (2018) 14:214–24. doi: 10.1038/nrneurol.2018.4
7. Ramdas S, Servais L. New treatments in spinal muscular atrophy: an overview of currently available data. *Expert Opin. Pharmacother.* (2020) 21:307–15. doi: 10.1080/14656566.2019.1704732
8. Mazzone ES, Mayhew A, Montes J, Ramsey D, Fanelli L, Young SD, et al. Revised upper limb module for spinal muscular atrophy: development of a new module. *Muscle Nerve.* (2017) 55:869–74. doi: 10.1002/mus.25430

9. O'Hagen JM, Glanzman AM, McDermott MP, Ryan PA, Flickinger J, Quigley J, et al. An expanded version of the Hammersmith functional motor scale for SMA II and III patients. *Neuromuscul. Disord.* (2007) 17:693–7. doi: 10.1016/j.nmd.2007.05.009
10. Dunaway Young S, Montes J, Kramer SS, Marra J, Salazar R, Cruz R, et al. Six-minute walk test is reliable and valid in spinal muscular atrophy. *Muscle Nerve.* (2016) 54:836–42. doi: 10.1002/mus.25120
11. Russman BS. Spinal muscular atrophy: clinical classification and disease heterogeneity. *J. Child Neurol.* (2007) 22:946–51. doi: 10.1177/0883073807305673
12. Arnold WD, Kassar D, Kissel JT. Spinal muscular atrophy: diagnosis and management in a new therapeutic era. *Muscle Nerve.* (2015) 51:157–67. doi: 10.1002/mus.24497
13. Swoboda KJ, Prior TW, Scott CB, McNaught TP, Wride MC, Reyna SP, et al. Natural history of denervation in SMA: relation to age, SMN2 copy number, and function. *Ann. Neurol.* (2005) 57:704–12. doi: 10.1002/ana.20473
14. Alves CRR, Zhang R, Johnstone AJ, Garner R, Nwe PH, Siranosian JJ, et al. Serum creatinine is a biomarker of progressive denervation in spinal muscular atrophy. *Neurology.* (2020) 94:e921–31. doi: 10.1212/WNL.0000000000008762
15. Carlier PG, Marty B, Scheidegger O, Loureiro De Sousa P, Baudin PY, Snezhko E, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging and spectroscopy as an outcome measure for clinical trials. *J. Neuromuscul. Dis.* (2016) 3:1–28. doi: 10.3233/JND-160145
16. Paoletti M, Pichiecchio A, Piccinelli SC, Tasca G, Berardinelli AL, Padovani A, et al. Advances in quantitative imaging of genetic and acquired myopathies: clinical applications and perspectives. *Front. Neurol.* (2019) 10:78. doi: 10.3389/fneur.2019.00078
17. Durmus H, Yilmaz R, Gulsen-Parman Y, Oflazer-Serdaroglu P, Cuttini M, Dursun MM, et al. Muscle magnetic resonance imaging in spinal muscular atrophy type 3: selective and progressive involvement. *Muscle Nerve.* (2017) 55:651–6. doi: 10.1002/mus.25385
18. Mercuri E, Talim B, Moghadaszadeh B, Petit N, Brockington M, Counsell S, et al. Clinical and imaging findings in six cases of congenital muscular dystrophy with rigid spine syndrome linked to chromosome 1p (RSMD1). *Neuromuscul. Disord.* (2002) 12:631–8. doi: 10.1016/S0960-8966(02)00023-8
19. Barp A, Carraro E, Albamonte E, Salmin F, Lunetta C, Comi G, et al. Muscle MRI in two SMA patients on nusinersen treatment: a two years follow-up. *J. Neurol. Sci.* (2020) 417:117067. doi: 10.1016/j.jns.2020.117067
20. Otto LAM, van der Pol WL, Schlaffke L, Wijngaarde CA, Stam M, Wadman RI, et al. Quantitative MRI of skeletal muscle in a cross-sectional cohort of patients with spinal muscular atrophy types 2 and 3. *NMR Biomed.* (2020) 33:e4357. doi: 10.1002/nbm.4357
21. Bakshi R, Dandamudi VSR, Neema M, De C, Bermel RA. Measurement of brain and spinal cord atrophy by magnetic resonance imaging as a tool to monitor multiple sclerosis. *J. Neuroimaging.* (2005) 15:30–45. doi: 10.1177/1051228405283901
22. Yiannakas MC, Mustafa AM, De Leener B, Kearney H, Tur C, Altmann DR, et al. Fully automated segmentation of the cervical cord from T1-weighted MRI using PropSeg: application to multiple sclerosis. *NeuroImage Clin.* (2016) 10:71–7. doi: 10.1016/j.nicl.2015.11.001
23. El Mendili MM, Querin G, Bede P, Pradat PF. Spinal cord imaging in amyotrophic lateral sclerosis: historical concepts—novel techniques. *Front. Neurol.* (2019) 10:350. doi: 10.3389/fneur.2019.00350
24. Wimmer T, Schreiber F, Hensiek N, Garz C, Kaufmann J, Machts J, et al. The upper cervical spinal cord in ALS assessed by cross-sectional and longitudinal 3T MRI. *Sci. Rep.* (2020) 10:1783. doi: 10.1038/s41598-020-58687-z
25. Prados F, Cardoso MJ, Yiannakas MC, Hoy LR, Tebaldi E, Kearney H, et al. Fully automated grey and white matter spinal cord segmentation. *Sci. Rep.* (2016) 6:36151. doi: 10.1038/srep36151
26. Yiannakas MC, Kearney H, Samson RS, Chard DT, Ciccarelli O, Miller DH, et al. Feasibility of grey matter and white matter segmentation of the upper cervical cord *in vivo*: a pilot study with application to magnetisation transfer measurements. *Neuroimage.* (2012) 63:1054–9. doi: 10.1016/j.neuroimage.2012.07.048
27. El Mendili MM, Lenglet T, Stojkovic T, Behin A, Guimarães-Costa R, Salachas F, et al. Cervical spinal cord atrophy profile in adult SMN1-linked SMA. *PLoS ONE.* (2016) 11:e0167886. doi: 10.1371/journal.pone.0167886
28. Querin G, El Mendili MM, Lenglet T, Behin A, Stojkovic T, Salachas F, et al. The spinal and cerebral profile of adult spinal-muscular atrophy: a multimodal imaging study. *NeuroImage Clin.* (2019) 21:101618. doi: 10.1016/j.nicl.2018.101618
29. Stam M, Haakma W, Kuster L, Froeling M, Philippens MEP, Bos C, et al. Magnetic resonance imaging of the cervical spinal cord in spinal muscular atrophy. *NeuroImage Clin.* (2019) 24:102002. doi: 10.1016/j.nicl.2019.102002
30. Michelson D, Ciafaloni E, Ashwal S, Lewis E, Narayanaswami P, Oskoui M, et al. Evidence in focus: nusinersen use in spinal muscular atrophy report of the guideline development, dissemination, and implementation subcommittee of the American academy of neurology. *Neurology.* (2018) 91:923–33. doi: 10.1212/WNL.0000000000006502
31. Alley S, Gilbert G, Wheeler-Kingshott CAMG, Samson RS, Grussu F, Martin A, et al. Consensus acquisition protocol for quantitative MRI of the cervical spinal cord at 3T. In: *Proceeding of the Joint Annual Meeting ISMRM-ESMRMB*. Paris (2018).
32. Blasche M, Riffel P, Lichy M. TimTX TrueShape and syngo ZOOMit technical and practical aspects (2012). p. 74–84.
33. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage.* (2006) 31:1116–28. doi: 10.1016/j.neuroimage.2006.01.015
34. Berglund J, Kullberg J. Three-dimensional water/fat separation and T2* estimation based on whole-image optimization—application in breathhold liver imaging at 1.5 T. *Magn. Reson. Med.* (2012) 67:1684–93. doi: 10.1002/mrm.23185
35. Smith DS, Berglund J, Kullberg J, Ahlström H, Avison MJ, Brian Welch E. Optimization of fat-water separation algorithm selection and options using image-based metrics with validation by ISMRM fat-water challenge datasets. In: *Proceedings of the International Society for Magnetic Resonance in Medicine*. Vol. 21 (2013).
36. Weigel M. Extended phase graphs: dephasing, RF pulses, and echoes - pure and simple. *J. Magn. Reson. Imaging.* (2015) 41:266–95. doi: 10.1002/jmri.24619
37. Marty B, Baudin P-Y, Reyngoudt H, Azzabou N, Araujo ECA, Carlier PG, et al. Simultaneous muscle water T2 and fat fraction mapping using transverse relaxometry with stimulated echo compensation. *NMR Biomed.* (2016) 29:431–43. doi: 10.1002/nbm.3459
38. Santini F, Deligianni X, Paoletti M, Weigel M, de Sousa PL, Bieri O, et al. A fast open-source implementation of water T2 with integrated fat fraction measurements from multi-echo spin-echo acquisitions. In: *Proceeding of MYOMRI*. (2019).
39. Brogna C, Cristiano L, Verdolotti T, Pichiecchio A, Cinnante C, Sansone V, et al. MRI patterns of muscle involvement in type 2 and 3 spinal muscular atrophy patients. *J. Neurol.* (2020) 267:898–912. doi: 10.1007/s00415-019-09646-w
40. De Leener B, Lévy S, Dupont SM, Fonov VS, Stikov N, Louis Collins D, et al. SCT: spinal cord toolbox, an open-source software for processing spinal cord MRI data. *Neuroimage.* (2017) 145:24–43. doi: 10.1016/j.neuroimage.2016.10.009
41. Gros C, De Leener B, Badji A, Maranzano J, Eden D, Dupont SM, et al. Automatic segmentation of the spinal cord and intramedullary multiple sclerosis lesions with convolutional neural networks. *Neuroimage.* (2019) 184:901–15. doi: 10.1016/j.neuroimage.2018.09.081
42. De Leener B, Fonov VS, Collins DL, Callot V, Stikov N, Cohen-Adad J. PAM50: unbiased multimodal template of the brainstem and spinal cord aligned with the ICBM152 space. *Neuroimage.* (2018) 165:170–9. doi: 10.1016/j.neuroimage.2017.10.041
43. Perone CS, Calabrese E, Cohen-Adad J. Spinal cord gray matter segmentation using deep dilated convolutions. *Sci. Rep.* (2018) 8:5966. doi: 10.1038/s41598-018-24304-3
44. Chang LC, Jones DK, Pierpaoli C. RESTORE: robust estimation of tensors by outlier rejection. *Magn. Reson. Med.* (2005) 53:1088–95. doi: 10.1002/mrm.20426
45. Helms G, Dathe H, Kallenberg K, Dechent P. High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and

T1 relaxation obtained from 3D FLASH MRI. *Magn. Reson. Med.* (2008) 60:1396–407. doi: 10.1002/mrm.21732
46. Kang PB, Krishnamoorthy KS, Jones RM, Shapiro FD, Darras BT. Atypical presentations of spinal muscular atrophy type III (Kugelberg-Welander disease). *Neuromuscul. Disord.* (2006) 16:492–4. doi: 10.1016/j.nmd.2006.05.004
47. Hagenacker T, Wurster CD, Günther R, Schreiber-Katz O, Osmanovic A, Petri S, et al. Nusinersen in adults with 5q spinal muscular atrophy: a non-interventional, multicentre, observational cohort study. *Lancet Neurol.* (2020) 19:317–25. doi: 10.1016/S1474-4422(20)30037-5
48. Maggi L, Bello L, Bonanno S, Govoni A, Caponnetto C, Passamano L, et al. Nusinersen safety and effects on motor function in adult spinal muscular atrophy type 2 and 3. *J. Neurol. Neurosurg. Psychiatry.* (2020) 91:1166–74. doi: 10.1136/jnnp-2020-323822
49. Bonati U, Holiga Š, Hellbach N, Risterucci C, Bergauer T, Tang W, et al. Longitudinal characterization of biomarkers for spinal muscular atrophy. *Ann. Clin. Transl. Neurol.* (2017) 4:292–304. doi: 10.1002/acn3.406
50. Chabanon A, Annoussamy M, Daron A, Pereon Y, Cances C, Vuillerot C, et al. Longitudinal data of the European prospective natural history study of patients with type 2 and 3 spinal muscular atrophy. *Neuromuscul. Disord.* (2017) 27:S134. doi: 10.1016/j.nmd.2017.06.154
51. Pichiecchio A, Rossi M, Cinnante C, Colafati GS, De Icco R, Parini R, et al. Muscle MRI of classic infantile pompe patients: fatty substitution and edema-like changes. *Muscle Nerve.* (2017) 55:841–8. doi: 10.1002/mus.25417
52. Kuru S, Sakai M, Konagaya M, Yoshida M, Hashizume Y, Saito K. An autopsy case of spinal muscular atrophy type III (Kugelberg-Welander disease). *Neuropathology.* (2009) 29:63–7. doi: 10.1111/j.1440-1789.2008.00910.x
53. Duval T, Lévy S, Stikov N, Campbell J, Mezer A, Witzel T, et al. g-ratio weighted imaging of the human spinal cord in *vivo*. *Neuroimage.* (2016) 145:11–23. doi: 10.1016/j.neuroimage.2016.09.018
54. Winder TR, Auer RN. Sensory neuron degeneration in familial Kugelberg-Welander disease. *Can. J. Neurol. Sci.* (1989) 16:67–70. doi: 10.1017/S0317167100028535
55. Araki S, Hayashi M, Tamagawa K, Saito M, Kato S, Komori T, et al. Neuropathological analysis in spinal muscular atrophy type II. *Acta Neuropathol.* (2003) 106:441–8. doi: 10.1007/s00401-003-0743-9
56. Paulson GW, Liss L, Sweeney PJ. Late onset spinal muscle atrophy - a sex linked variant of Kugelberg-Welander. *Acta Neurol. Scand.* (1980) 61:49–55. doi: 10.1111/j.1600-0404.1980.tb02995.x
57. Sheng L, Rigo F, Bennett CF, Krainer AR, Hua Y. Comparison of the efficacy of MOE and PMO modifications of systemic antisense oligonucleotides in a severe SMA mouse model. *Nucleic Acids Res.* (2020) 48:2853–65. doi: 10.1093/nar/gkaa126
58. Berciano MT, Puente-Bedia A, Medina-Samamé A, Rodríguez-Rey JC, Calderó J, Lafarga M, et al. Nusinersen ameliorates motor function and prevents motoneuron Cajal body disassembly and abnormal poly(A) RNA distribution in a SMA mouse model. *Sci. Rep.* (2020) 10:10738. doi: 10.1038/s41598-020-67569-3
59. Pena E, Berciano MT, Fernandez R, Ojeda JL, Lafarga M. Neuronal body size correlates with the number of nucleoli and Cajal bodies, and with the organization of the splicing machinery in rat trigeminal ganglion neurons. *J. Comp. Neurol.* (2001) 430:250–63. doi: 10.1002/1096-9861(20010205)430:2<250::AID-CNE1029>3.0.CO;2-L

Conflict of Interest: AP is a member of the advisory board and consultant for G-enzyme Sanofi.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

CASE REPORT
published: 09 April 2021
doi: 10.3389/fneur.2021.644893

Case Report: Neurogenic Thoracic Outlet Syndrome Without Electrophysiologic Abnormality

*Sun Woong Kim and Duk Hyun Sung**

Department of Physical and Rehabilitation Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, South Korea

Neurogenic thoracic outlet syndrome (N-TOS) is a chronic compressive brachial plexopathy that involves the C8, T1 roots, and/or lower trunk. Medial antebrachial cutaneous (MABC) nerve conduction study (NCS) abnormality is reportedly one of the most sensitive findings among the features of N-TOS. The aim of the present study was to report clinical features, imaging findings, treatment, and prognoses of two N-TOS patients with no abnormalities in electrophysiological studies. Both patients presented with paresthesia of unilateral arm, and examination revealed no neurologic deficits. Electrophysiologic studies including MABC NCS were normal. Computed tomography (CT) angiography and brachial plexus magnetic resonance imaging (MRI) of the patients showed compression and displacement of the neurovascular bundle in the thoracic outlet by causative structures. Due to their sensory symptoms and CT angiography and brachial plexus MRI findings, after excluding other diseases, we diagnosed them with N-TOS. With the development of imaging techniques, more patients presenting with clinical features of lower trunk brachial plexopathy and anomalous structures compressing the neurovascular bundle on imaging studies can be diagnosed with N-TOS, even if electrophysiologic studies including MABC NCS do not show abnormalities.

Keywords: neurogenic thoracic outlet syndrome, CT angiography, brachial plexus magnetic resonance imaging, medial antebrachial cutaneous nerve, case report

OPEN ACCESS

Edited by:
Massimiliano Filosto, University of Brescia; NeMO-Brescia Clinical Center for Neuromuscular Diseases, Italy

Reviewed by:
Stefano Carlo Previtali, San Raffaele Scientific Institute (IRCCS), Italy
Richard Sanders, University of Colorado, United States

***Correspondence:**
Duk Hyun Sung
yays.sung@samsung.com
orcid.org/0000-0002-8261-7199

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: *22 December 2020*
Accepted: *17 March 2021*
Published: *09 April 2021*

Citation:
Kim SW and Sung DH (2021) Case Report: Neurogenic Thoracic Outlet Syndrome Without Electrophysiologic Abnormality. Front. Neurol. 12:644893. doi: 10.3389/fneur.2021.644893

INTRODUCTION

Neurogenic thoracic outlet syndrome (N-TOS) is a chronic compressive brachial plexopathy that involves the C8, T1 roots, and/or lower trunk (1). In N-TOS, the brachial plexus is pressed anteriorly and upwardly by the cervical rib or fibrous band, arising elongated transverse process of C7. The disease mainly invades through the T1 root because it is located at the lowermost side of the brachial plexus. As a result, the median-innervated intrinsic hand muscles receiving T1-dominant innervation are usually involved (2). The sensory symptom also occurs mainly in the medial side of the forearm, which is medial antebrachial cutaneous (MABC) nerve territory and T1 dermatome. Within the same mechanism, MABC sensory nerve conduction study (NCS) abnormalities are reportedly the most sensitive finding among the electrophysiological features of N-TOS (3, 4). However, we recently encountered two patients with no abnormalities in electrophysiological studies, including MABC NCS, although their clinical presentations and imaging findings were compatible with N-TOS. Here, we report their clinical features, imaging findings, treatment, and prognoses.

CASE DESCRIPTION

Case 1

A 49-year-old woman visited our locomotor pain clinic with a 4-year history of paresthesia in her left medial arm, medial forearm, and fourth and fifth fingers. Her symptoms were aggravated in the supine and left lateral decubitus positions and on rotating her head to the left side, and often woke her during sleep. The paresthesia was also aggravated when her left supraclavicular fossa area was compressed. Epidural steroid injection in the cervical spine at her local community hospital did not improve her paresthesia. Neurological examination showed no motor weaknesses or sensory deficits. Percussion of the left supraclavicular fossa provoked paresthesia of her left medial arm, medial forearm, and fourth and fifth fingers. An Adson's test was positive on her left side, showing the loss of her radial pulse. The test evaluate the reproduction of symptoms or loss of radial pulse by extending the neck and rotating the head to the symptomatic side while holding a deep inspiration. Spurling and upper extremity tension tests were negative. Sensory NCS was done in the bilateral median, ulnar and MABC nerves with the antidromic method. Median and ulnar motor NCS was also done on both sides, recorded in abductor pollicis brevis and abductor digiti minimi muscles, respectively. Age-stratified reference values derived from our electrodiagnostic laboratories were used for absolute criteria of amplitude abnormality. Needle EMG was done in limb muscles, especially the C8-T1 myotome, and cervical paraspinal muscles. These electrophysiological studies did not show any abnormalities. Computed tomography (CT) angiography of her upper extremities performed in 180° shoulder abduction showed the left cervical rib arising from the C7 vertebra and mild stenosis of the left subclavian artery near the lateral tip of the cervical rib, suggesting arterial thoracic outlet syndrome (A-TOS) (**Figures 1A–C**). The test showed no dissection, aneurysm, or thrombosis. Brachial plexus magnetic resonance imaging (MRI) revealed upward and anterior angulation and displacement of the left lower trunk or C8/T1 extraforaminal nerve roots of the brachial plexus, presumably caused by the anterior tip of the cervical rib (**Figures 1D,E**). Swelling or abnormal signal intensity of the brachial plexus were not definite. Evidence of C8 or T1 cervical radiculopathy because of neural foraminal stenosis or herniated intervertebral disc was not observed in cervical spine MRI. The results of the imaging tests indicated that the cervical rib was compressing the lower trunk/roots of the brachial plexus and the subclavian artery upwardly and anteriorly. With the sensory symptoms suggesting a lesion of the lower trunk/roots of the brachial plexus and cervical rib, along with findings of the CT angiography and brachial plexus MRI, and by excluding other diseases, we diagnosed the patient with N-TOS despite normal electrophysiological study findings. We decided to manage her conservatively because there were no ischemic symptoms/signs and focal neurologic deficits in her left upper extremities. We prescribed neuropathic pain medication (gabapentin 100 mg twice a day), which helped relieve the patient's paresthesia. The patient was trained to avoid behaviors that may cause neurovascular bundle compression, and instructed to observe if there was any weakness or numbness. We also planned to conduct imaging studies to follow the status of the subclavian artery at periodic intervals. At the last follow-up 12 months after her initial visit, although better than at initial visit, the paresthesia while lying on her left side persisted. She remained positive for Tinel's sign at the supraclavicular fossa, although weakness, atrophy, or sensory loss was not detected.

Case 2

A 31-year-old man with a 4-month history of paresthesia on his right medial forearm and three ulnar digits was referred to our locomotor pain clinic. His paresthesia would get aggravated when lying on the right side and raising the right arm overhead for tasks, such as, when washing his hair. Neurological examination showed normal muscle strength and sensory function of his right upper extremities. No atrophy was noted in the intrinsic muscles of his right hand. Tinel's test was negative for his right supraclavicular fossa. At 90° abduction and external rotation of the shoulder (Roos test posture), the pain in his right palm and forearm worsened, and his right palm turned pale. An electrophysiological study performed at his local community hospital was reviewed by the authors. When the reference value of our institute was applied, NCS including MABC nerve did not show any abnormality. Needle EMG result was also normal. The patient's right first rib was anomalous (upwardly convex and anterior tip fused with the second rib) in cervical spine plain radiograph (**Figure 2A**). CT angiography of his upper extremities showed stenosis of the right subclavian artery overlying the anomalous first rib, suggesting A-TOS. The subclavian artery was dilated distal to the stenotic site. Segmental thromboembolic stenosis of the right brachial artery was also observed (**Figures 2B–D**). Brachial plexus MRI revealed high-riding T1 extraforaminal nerve root or lower trunk compared with that of the left side and anterior angulation of the right lower trunk by the anomalous right first rib (**Figures 2E,F**). Swelling or abnormal signal intensity of the brachial plexus was not definite. There was no evidence of lower cervical radiculopathy on MRI. Based on his sensory symptoms, anomalous right first rib, and CT angiography and brachial plexus MRI findings, and after excluding other diseases, we diagnosed the patient with N- and A-TOS despite normal electrophysiological study findings. The patient underwent surgical treatment (first rib resection, scalenectomy, and brachio-ulnar bypass with great saphenous vein graft) to resolve the neurovascular bundle compression and thrombosis of the brachial artery. The patient's paresthesia disappeared immediately after the operation.

DISCUSSION

TOS includes variable neurovascular symptoms and signs caused by compression of neurovascular bundle in the thoracic outlet. TOS is divided into three types depending on the compressed structures: arterial, venous, and neurogenic. Technical improvements in MRI is facilitating detection of brachial plexus compression in patients with N-TOS. Treatment options for N-TOS include surgery, physical therapy, behavioral

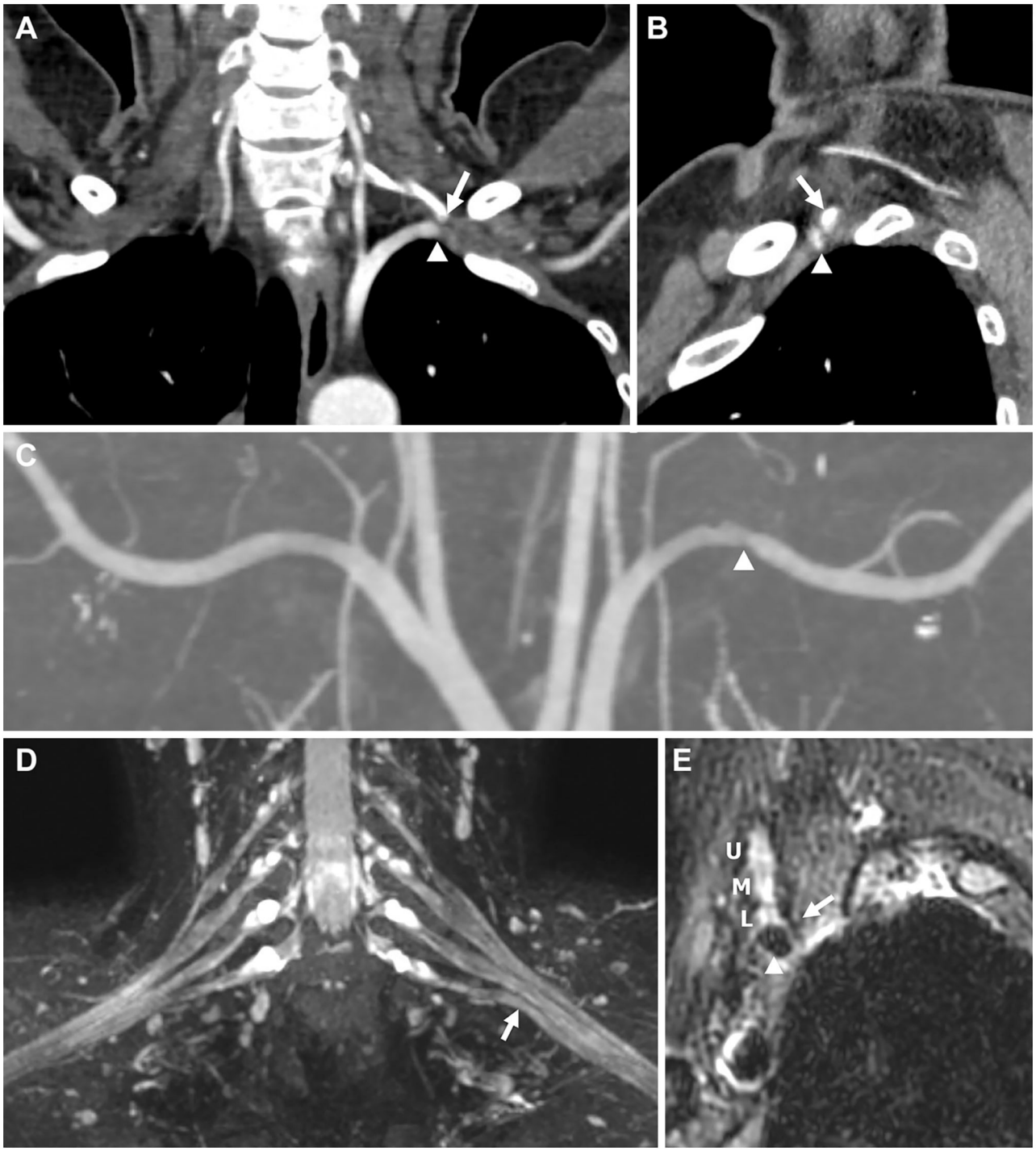


FIGURE 1 | Imaging findings for Patient 1. **(A,B)** Computed tomography (CT) angiography performed with 180° shoulder abduction. **(C–E)** Brachial plexus magnetic resonance imaging (MRI). **(A)** Coronal image showing the tip of the left cervical rib arising from the C7 vertebra (arrow) abutting the left subclavian artery (arrowhead). **(B)** Sagittal image showing the tip (arrow) compressing the subclavian artery (arrowhead) in the anterior direction, causing stenosis. **(C)** Coronal maximum intensity projection (MIP) image showing stenosis of the left subclavian artery (arrowhead). **(D)** Coronal T2-weighted fat suppression MIP image showing upward angulation of the lower trunk (arrow), presumably caused by the anterior tip of the cervical rib, although not shown in this image. Swelling or abnormal signal intensity of the brachial plexus was not definite. **(E)** Oblique sagittal T2-weighted fat suppression image showing the spatial associations of the anterior tip of the cervical rib (arrow), brachial plexus (U, upper trunk; M, middle trunk; L, lower trunk), and subclavian artery (arrowhead).

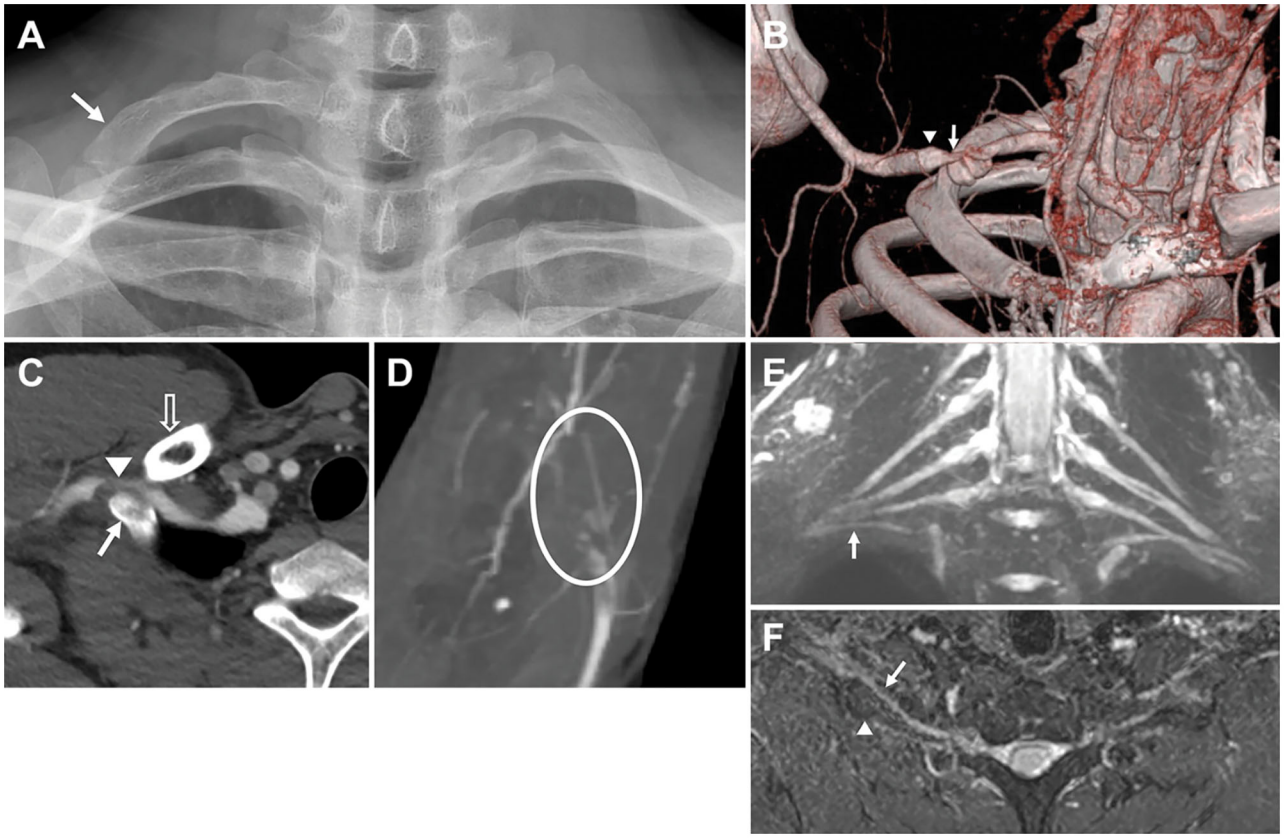


FIGURE 2 | Imaging findings for Patient 2. **(B–D)** Computed tomography (CT) angiography performed with 180° shoulder abduction. **(E,F)** Brachial plexus magnetic resonance imaging (MRI). **(A)** Plain radiograph of the cervical spine showing upwardly convex curvature and an anomaly of the right first rib (arrow). **(B)** When viewed from the front, a reconstruction image using volume rendering shows stenosis (arrow) and post-stenotic dilatation (arrowhead) of the right subclavian artery. **(C)** Axial image showing the spatial relationships of the anomalous right first rib (arrow) articulating with the second rib, narrowed subclavian artery (arrowhead), and clavicle (voided arrow). **(D)** Maximum intensity projection (MIP) image showing thromboembolic occlusion of the right brachial artery at the elbow level (circle). **(E)** Coronal T2-weighted fat suppression MIP image showing a high-riding right T1 extraforaminal root and lower trunk (arrow) compared to those on the opposite side. Abnormal signal intensity of the brachial plexus was not definite. **(F)** Axial T2-weighted fat suppression MIP image showing anterior angulation of the right lower trunk (arrow) by the anomalous right first rib (arrowhead).

modification, and pain management. It is the symptoms and physical examination findings that determines the type of treatment, rather than the presence of rib abnormalities or the results of electrophysiologic studies. Usually, initial management of N-TOS consists of conservative treatments. For patients with refractory symptoms or profound focal neurologic deficit, consideration of surgical management is needed. Optimal conservative treatment for N-TOS is still controversial. For patients with A- or V-TOS, initial intervention is most often surgical because physical therapy is not helpful (5).

Both patients described in this report showed typical neurologic symptoms of N-TOS, namely, paresthesia of the medial forearm and ulnar hand, and structural anomalies of the thoracic outlet (cervical rib and anomalous first rib). Diseases such as C8 or T1 radiculopathy, carpal tunnel syndrome, and ulnar neuropathy, which require differentiation from N-TOS, were excluded on the basis of physical examination, MRI, and electrophysiological study. It is reasonable to infer that compression of the neurovascular bundle by the structural anomalies causes neurologic symptoms. However, no neurologic deficit was observed in the neurologic examinations, and no abnormalities were observed in the electrophysiological studies. The traditional concept indicates that, because N-TOS is a slow progressing disease, patients seek medical care when overt deficits, such as weakness or atrophy of the T1 myotome muscles, occur (6). As a result, sensory abnormalities are less pronounced than weakness or atrophy (7). However, a recent study has shown that this is not the case, because more than half of the patients reported sensory symptoms as their chief complaint (3). In addition, about one-third of patients had no motor symptoms at initial presentation. Both patients in the present study also showed only sensory symptoms. Because sensory fibers are more sensitive to compression than motor fibers, it is reasonable that the sensory symptom appears early.

In electrophysiological studies, N-TOS typically reveals a characteristic "median-motor, MABC and ulnar-sensory" pattern

(4, 8). The MABC is more involved than the ulnar nerve in sensory NCS because of the T1 predominance, as MABC nerve territory corresponds to the T1 dermatome and the ulnar nerve territory to the C8 dermatome (4, 9). Seror et al. (10) suggested that the MABC NCS could be used for the detection of mild lower brachial plexus lesions because they observed abnormal MABC nerve findings in all patients with unilateral, atypical pain and paresthesia of the upper limbs. Kim et al. (3) also observed MABC nerve abnormalities in all 13 patients with N-TOS and suggested that MABC NCS could be used as a screening test in patients with N-TOS because of its high sensitivity.

However, unlike previous reports, this report showed that MABC NCS is not a 100% sensitive test. In a literature review, Tsao et al. (4) described the electrodiagnostic findings of patients with N-TOS with surgical verification. Among 19 patients who underwent MABC NCS, one patient showed no abnormality. In a study of six surgically confirmed patients with N-TOS, Le Forestier et al. (8) reported one patient with normal MABC NCS. The possible explanation for the negative MABC NCS finding of N-TOS is the nature of compressive peripheral neuropathy. The pathophysiology of compressive neuropathy, such as carpal tunnel syndrome (CTS) or cubital tunnel syndrome (CuTs), is typically demyelination and may be associated with secondary axonal loss (11). Therefore, prolonged latency, reduced conduction velocity, or conduction block are usually observed initially, before amplitude reduction appears. In the case of MABC NCS, a latency delay cannot be observed because it is stimulated from the distal side of the lesion, unlike the median NCS in CTS. Therefore, patients with N-TOS in the very early stage of compressive neuropathy, before an axonal loss occurs, may not present electrophysiological abnormalities. Additionally, there are some reports that the MABC nerve is derived from both C8 and T1 dorsal root ganglia rather than T1 alone (12). In N-TOS, the T1 root is thought to receive initial compressive insult anatomically. Because of the dual supply of the MABC nerve, T1 root lesions alone may not cause abnormalities or a significant decrease of amplitude by more than 50% compared with the opposite side as the C8 root is intact.

Reports on the image features of N-TOS are scanty, mostly at the case report level, and there are no established criteria. Previous study suggested four imaging criteria for N-TOS; namely, upwardly-pushed neural structures in brachial plexus MRI, increased signal intensity of the lower roots/trunk in brachial plexus MRI, focal stenosis of the subclavian artery in CT angiography, and high-mounted subclavian artery in CT angiography (3). Because these criteria showed relatively low sensitivity (22–70%), the authors of the study recommended that imaging studies play ancillary roles in diagnosis. Baumer et al. (7) reported that only 7 of 30 patients showed abnormalities in brachial plexus MRI, which also showed low sensitivity. However, abnormalities in the imaging studies can play a definite role in diagnosis of N-TOS when there is no abnormality in electrophysiologic study. In this study, our two patients were diagnosed with N-TOS based only on the findings of imaging studies despite negative results for electrophysiologic studies including MABC NCS. Because the compressive neuropathy may progress over time and neurologic deficits or electrophysiologic abnormalities may appear later, careful monitoring of neuropathic symptoms/signs is required for patients in very early stages, especially those who had not undergone surgical decompression for causative structural abnormalities.

In conclusion, patients presenting with clinical features of lower trunk brachial plexopathy and anomalous structures compressing the neurovascular bundle on imaging studies can be diagnosed with N-TOS. The development of imaging techniques may reveal more such cases. Therefore, for these patients, brachial plexus MRI and CT angiography may be helpful to diagnose N-TOS and accompanying vascular TOS, even if electrophysiologic studies including MABC NCS do not show abnormalities.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional board registry of Samsung Medical Center (registry number: 2020-08-010). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

SK and DS: conceptualization and design of the work, acquisition, analysis, or interpretation of data, and drafting the work or revising it critically for important intellectual content. All authors have read and approved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.644893/full#supplementary-material

REFERENCES

1. Ferrante MA. The thoracic outlet syndromes. *Muscle Nerve.* (2012) 45:780–95. doi: 10.1002/mus.23235

2. Chiba T, Konoeda F, Higashihara M, Kamiya H, Oishi C, Hatanaka Y, et al. C8 and T1 innervation of forearm muscles. *Clin Neurophysiol.* (2015) 126:837–42. doi: 10.1016/j.clinph.2014.07.031

3. Kim SW, Jeong JS, Kim BJ, Choe YH, Yoon YC, Sung DH. Clinical, electrodiagnostic and imaging features of true neurogenic thoracic outlet syndrome: experience at a tertiary referral center. *J Neurol Sci.* (2019) 404:115–23. doi: 10.1016/j.jns.2019.07.024
4. Tsao BE, Ferrante MA, Wilbourn AJ, Shields RW. Electrodiagnostic features of true neurogenic thoracic outlet syndrome. *Muscle Nerve.* (2014) 49:724–7. doi: 10.1002/mus.24066
5. Jones MR, Prabhakar A, Viswanath O, Urits I, Green JB, Kendrick JB, et al. Thoracic outlet syndrome: a comprehensive review of pathophysiology, diagnosis, and treatment. *Pain Ther.* (2019) 8:5–18. doi: 10.1007/s40122-019-0124-2
6. Ferrante MA, Ferrante ND. The thoracic outlet syndromes: Part 1. Overview of the thoracic outlet syndromes and review of true neurogenic thoracic outlet syndrome. *Muscle Nerve.* (2017) 55:782–93. doi: 10.1002/mus.25536
7. Baumer P, Kele H, Kretschmer T, Koenig R, Pedro M, Bendszus M, et al. Thoracic outlet syndrome in 3T MR neurography-fibrous bands causing discernible lesions of the lower brachial plexus. *Eur Radiol.* (2014) 24:756–61. doi: 10.1007/s00330-013-3060-2
8. Le Forestier N, Moulonguet A, Maisonobe T, Leger JM, Bouche P. True neurogenic thoracic outlet syndrome: electrophysiological diagnosis in six cases. *Muscle Nerve.* (1998) 21:1129–34. doi: 10.1002/(SICI)1097-4598(199809)21:9<1129::AID-MUS3>3.0.CO;2-9
9. Ferrante MA, Wilbourn AJ. The utility of various sensory nerve conduction responses in assessing brachial plexopathies. *Muscle Nerve.* (1995) 18:879–89. doi: 10.1002/mus.880180813
10. Seror P. Medial antebrachial cutaneous nerve conduction study, a new tool to demonstrate mild lower brachial plexus lesions. A report of 16 cases. *Clin Neurophysiol.* (2004) 115:2316–22. doi: 10.1016/j.clinph.2004.04.023
11. Rempel DM, Diao E. Entrapment neuropathies: pathophysiology and pathogenesis. *J Electromyogr Kinesiol.* (2004) 14:71–5. doi: 10.1016/j.jelekin.2003.09.009
12. Stylianos K, Konstantinos G, Pavlos P, Aliki F. Brachial branches of the medial antebrachial cutaneous nerve: a case report with its clinical significance and a short review of the literature. *J Neurosci Rural Pract.* (2016) 7:443–6. doi: 10.4103/0976-3147.182772

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 16 April 2021
doi: 10.3389/fneur.2021.630435

Quantification and Monitoring of the Effect of Botulinum Toxin A on Paretic Calf Muscles of Children With Cerebral Palsy With MRI: A Preliminary Study

Claudia Weidensteiner[1,2], Philipp Madoerin[1], Xeni Deligianni[1,2], Tanja Haas[1], Oliver Bieri[1,2], Tugba Akinci D'Antonoli[3,4], Katrin Bracht-Schweizer[5], Jacqueline Romkes[5], Enrico De Pieri[2,5], Francesco Santini[1,2], Erich Rutz[6,7], Reinald Brunner[5,8] and Meritxell Garcia[9]*

[1] Division of Radiological Physics, Department of Radiology, University Hospital of Basel, Basel, Switzerland, [2] Department of Biomedical Engineering, University of Basel, Allschwil, Switzerland, [3] Department of Radiology, University Hospital of Basel, Basel, Switzerland, [4] Department of Radiology, University Children's Hospital Basel, Basel, Switzerland, [5] Laboratory for Movement Analysis, University Children's Hospital Basel, Basel, Switzerland, [6] Murdoch Children's Research Insitute, The University of Melbourne, Pediatric Orthopedic Department, The Royal Children's Hospital, Parkville, VIC, Australia, [7] Faculty of Medicine, The University of Basel, Basel, Switzerland, [8] Department of Orthopedic Surgery, University Children's Hospital Basel, Basel, Switzerland, [9] Department of Radiology, Division of Neuroradiology, University Hospital of Basel, Basel, Switzerland

OPEN ACCESS

Edited by:
Aleksandar Beric, New York University, United States

Reviewed by:
Matthias Klotz, Kepler University Hospital GmbH, Austria
Anja Van Campenhout, KU Leuven, Belgium

***Correspondence:**
Claudia Weidensteiner
claudia.weidensteiner@unibas.ch

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: *17 November 2020*
Accepted: *01 March 2021*
Published: *16 April 2021*

Citation:
Weidensteiner C, Madoerin P, Deligianni X, Haas T, Bieri O, Akinci D'Antonoli T, Bracht-Schweizer K, Romkes J, De Pieri E, Santini F, Rutz E, Brunner R and Garcia M (2021) Quantification and Monitoring of the Effect of Botulinum Toxin A on Paretic Calf Muscles of Children With Cerebral Palsy With MRI: A Preliminary Study. Front. Neurol. 12:630435. doi: 10.3389/fneur.2021.630435

Background: Muscles from patients with cerebral palsy (CP) are often spastic and form contractures that limit the range of motion. Injections of botulinum toxin A (BTX) into the calf muscles are an important treatment for functional equinus; however, improvement in gait function is not always achieved. BTX is also used to test muscle weakening for risk evaluation of muscle lengthening surgery. Our aim was to assess the effect of BTX over time on calf muscle properties in pediatric CP patients with MRI.

Material and Methods: Six toe-walking CP patients (mean age 11.6 years) with indication for lengthening surgery were prospectively enrolled and received BTX injections into the gastrocnemius and soleus muscles. MRI scans at 3T of the lower legs and clinical examinations were performed pre-BTX, 6 weeks (6w), and 12 weeks (12w) post-BTX. A fat-suppressed 2D multi-spin-echo sequence was used to acquire T_2 maps and for segmentation. Fat fraction maps were calculated from 3D multi-echo Dixon images. Diffusion tensor imaging (DTI) with a 2D echo-planar imaging (EPI) sequence yielded maps of the mean apparent diffusion coefficient (ADC) and of the fractional anisotropy (FA). Hyperintense regions of interest (ROIs) on the T_2-weighted (T_2w) images at 6w were segmented in treated muscles. Mean values of T_2, fat fraction, ADC, and FA were calculated in hyperintense ROIs and in reference ROIs in non-treated muscles.

Results: Hyperintensity on T_2w scans and increased T_2 (group mean ± standard deviation: 35 ± 1 ms pre-BTX, 45 ± 2 ms at 6w, and 44 ± 2 ms at 12w) were observed in all patients at the injection sites. The T_2 increase was spatially limited to parts of the injected muscles. FA increased (0.30 ± 0.03 pre-BTX, 0.34 ± 0.02 at 6w, and 0.36 ± 0.03 at 12w) while ADC did not change in hyperintense ROIs, indicating a

BTX-induced increase in extracellular space and a simultaneous decrease of muscle fiber diameter. Fat fraction showed a trend for increase at 12w. Mean values in reference ROIs remained unchanged.

Conclusion: MRI showed limited spatial distribution of the BTX-induced effects in pediatric CP patients. It could be a promising non-invasive tool for future studies to test BTX treatment protocols.

Keywords: cerebral palsy, MRI, botulinum toxin A, T_2, diffusion, fat fraction, calf muscles, pediatric

INTRODUCTION

Cerebral palsy (CP) is a sensorimotor dysfunction caused by damage to the developing brain and is the most common cause of lifelong motor disability in children (1), affecting movement and muscle coordination. Muscles from CP patients are often spastic and form contractures that limit the range of motion (RoM) and joint function. Botulinum toxin A (BTX) has been established as an important treatment modality, particularly for the management of spasticity (2) to reduce inadequate muscle activity in CP patients. BTX is injected locally into the affected calf muscles to control functional equinus and spasticity and thus improve gait function. In general, equinus treatment with BTX is moderately effective, especially in younger children under 6 years of age, with the clinical effect lasting for 3–6 months (3). There is no uniform BTX treatment strategy in CP; instead, there are several recommendations and guidelines about total doses, injection volume, and injection sites based on expert opinions and small clinical trials (2). However, it is unknown to what extent BTX is distributed into the muscle tissue and if it affects the muscle belly as a whole or only partially. BTX blocks the neurotransmission at the motor end plates, which are located in a specific zone within the muscle belly (4). Depending on the injection technique and on the subsequent diffusion of the drug, the complete zone of motor end plates or only a part of it is reached and affected by BTX. This leads to a variation in treatment response (4). Also, improvement in gait function is not always achieved with BTX treatment, especially in children over 6 years of age who may need a surgical lengthening of the calf muscles as an effective treatment for equinus (3). Thus, there is a need to study and monitor the effect on injected muscles.

There are various methods based on magnetic resonance imaging (MRI) that offer insight into muscle tissue changes. An increase in water content, i.e., muscle edema, leads to an increase in the MR transverse relaxation time T_2 value that manifests as hyperintense areas on T_2-weighted (T_2w) images. Muscle T_2w MR images and mainly quantitative T_2 relaxation values of water can assess inflammation, e.g., in thigh muscles in dermatomyositis (5, 6), and edema, e.g., in calf muscles in Duchenne dystrophy (7). The measurement of fat infiltration and replacement in muscles monitors the extent and severity of muscle destruction, especially in chronic diseases like Duchenne muscular dystrophy (8) or diabetic neuropathy (9). Diffusion-weighted imaging (DWI) and diffusion tensor imaging (DTI) are methods to investigate the tissue microstructural integrity since cell membranes and other structures restrict the diffusion of water molecules. In DTI, a diffusion tensor describing the diffusivity in different directions is generated from a series of diffusion-weighted images. From the eigenvalues of these tensors, the mean apparent diffusion coefficient (ADC; also called mean diffusivity MD in DTI) and the fractional anisotropy (FA)—a measure for diffusion anisotropy ranging from 0 (completely isotropic) to 1 (completely anisotropic)—can be calculated (10). These parameters vary with architectural changes occurring in muscles, e.g., in calf muscle injury (11) or in the lower leg of patients with muscle dystrophy (12).

Our aim was to assess the effect of BTX on muscle properties in pediatric CP patients with MRI over 12 weeks after the injections. We were interested in quantifying the following potential muscle tissue changes using MRI: the T_2 of muscle water, diffusion of muscle water, and fat content in the BTX-treated muscles. Results from this pilot study may help to better understand the physiological and therapeutic effects of BTX on muscles in patients with CP.

MATERIALS AND METHODS

Participants

Six ambulatory patients diagnosed with CP (five boys and one girl, mean age 11.6 years, range 9.8–12.8 years, three bilateral spastic CP functionally diparetic and three unilateral spastic CP functionally hemiparetic) were included and prospectively monitored between 2018 and 2020. All patients were toe-walking and were scheduled for muscle lengthening surgery with a preoperative BTX test injection in the gastrocnemius muscles to investigate muscle weakening, which can be a side effect of the muscle lengthening surgery (13). Exclusion criteria were previous surgeries on the affected limb(s), claustrophobia, and difficulties in following instructions in the scanner. Full participant characteristics can be found in **Table 1**. Functional mobility level was classified as Gross Motor Function Classification System level (GMFCS) I ($n = 4$), II ($n = 1$), and III ($n = 1$) (14). Prior to participation, informed written consent was obtained from the participants' parents and additionally from those participants age 12 years or older. The study was approved by the local ethics committee.

Treatment

BTX (onabotulinum toxin A, Botox, Allergan plc, Dublin, Ireland) was injected under ultrasound guidance in the gastrocnemius lateralis (GL), medialis (GM), and soleus (S) muscles (dose 50–100 units per muscle, one to two injection

TABLE 1 | Patient characteristics, treatment, and results of the clinical examination pre-BTX, 6 weeks (6w), and 12 weeks (12w) post-BTX showing lower limb spasticity, passive range of motion (RoM), and manual muscle testing (MMT).

Patient number	Patient characteristics						Treatment				Spasticity (Modified Ashworth scale)			Passive RoM (deg.)		MMT (Medical Research Council scale)	
											Pre/6w/12w			Pre/6w/12w		Pre/6w/12w	
	Age range pre-BTX (years)	Height pre-BTX (cm)	Body weight pre-BTX (kg)	BMI percentile pre-BTX	CP uni/bi	GMFCS	Side	BTX treatment naïve	BTX dose	BTX dose per body weight (U/kg)	PF (at 90° KF)	PF (at KE)	KF	DF (at KE)	KE (at HE)	PF	KF
1	12–14	146.5	32.7	5	uni	I	r	n	GM, GL, S: 50 U diluted in 1 ml at 1 injection site each	4.6	1+/1+/1	1+/1/1+	0/1/1	−10/−5/−10	0/0/0	2+/2+/2+	5/4/4
2	10–12	151.0	53.8	95	uni	I	r	n	GM, GL, S: 50 U diluted in 1 ml at 1 injection site each	2.8	0/1/1	1/1/0	0/0/1+	−20/−20/−30	−15/−20/−10	2+/2+/2+	4/4/4
3	10–12	146.0	34.4	23	uni	I	r	n	GM, GL, S: 50 U diluted in 1 ml at 1 injection site each	4.4	1+/1/1	1/1/1	0/0/0	15/10/5	0/0/0	2+/2+/3	5/5/5
4	10–12	137.0	32.6	49	bi	I	l	y	GM, GL, S: 50 U diluted in 1 ml at 1 injection site each	4.6	1/0/1	1/0/0	0/0/0	−10/−15/−15	−20/−20/−15	2+/2+/2+	4/4/4
5	10–12	135.0	26.6	3	bi	II	l	n	GM, GL, S: 75 U diluted in 1.5 ml distributed at 2 injection sites each	8.5	1+/2/1	4/1/1+	2/1/1	−30/−15/−15	−5/0/−10	3/2/2+	4/3+/3
6	8–10	127.5	22.7	4	bi	III	l	n	GM, GL: 100 U diluted in 2 ml distributed at 2 injection sites each	8.8	1/1+/1+	1/1+/0	1/1+/1	−5/−15/−15	−10/−15/−10	3+/2+/2+	3/3+/3+

BMI, body mass index; GMFCS, Gross Motor Function Classification System; GM, m. gastrocnemius medialis; GL, m. gastrocnemius lateralis; S, m. soleus; U, units; BTX, botulinum toxin A; CP, cerebral palsy; uni, unilateral; bi, bilateral; PF, plantarflexor muscles; KF, knee flexor muscles; DF, dorsiflexion; KE, knee extension; HE, hip extension.

sites per muscle, dilution: 50 units diluted to 1 ml). The injection sites were at a horizontal line separating the proximal 1/4 and the distal 3/4 of the muscle belly for the GM and GL and at a horizontal line separating the proximal 1/3 and the distal 2/3 of the muscle belly for the S. The dose was chosen according to the current state of the art recommendations (2, 15): the maximal dose per injection site was 50 units, and the dose per kg body weight was below 20 units/kg (**Table 1**). Target muscles were determined according to a prior gait analysis. The patients with bilateral paresis were injected in the more severely affected leg only. The injection was not intended to improve the patients' clinical symptoms but to test their reaction on weakness of this muscle for a risk evaluation of the planned lengthening surgery (13).

Study Protocol

MRI scans and routine clinical examinations were performed pre-BTX injection, as well as 6 weeks (6w) and 12 weeks (12w) post-BTX injection. Gait analysis [as described in Rutz et al. (13)] was performed on the same time points, but it is not reported here because it is out of the scope of this paper. The maximum BTX effect was expected 6w post-BTX (16). The second MRI scan was within 3 months after the first one.

Clinical Examination

Besides general parameters such as weight, height, and body mass index [BMI (17)], the passive RoM of the knee and ankle joints [passive RoM, using a goniometer (18)] and lower limb spasticity [Modified Ashworth Scale (19)] were measured. Manual muscle testing [MMT; Medical Research Council scale (20)] was performed in the legs to assess muscle strength.

MR Image Acquisition and Analysis

MRI exams were performed at a whole-body clinical scanner with 3T field strength (Siemens Prisma, Siemens Healthineers, Erlangen, Germany). The patients were positioned feet first supine on the MRI patient table, and the more severely affected leg was restrained with straps at a comfortable resting angle, with the knee in maximal extension. A Siemens 18-element-body array coil was placed on the lower legs. For T_2 mapping, a 2D multi-spin-echo sequence was used: 32 axial slices, voxel size 1.0 × 1.0 × 3.0 mm^3, reconstructed matrix size = 320 × 324 × 32, repetition time (TR) = 4.3 s, echo time (TE) = 11–115 ms, fat-saturated, acceleration factor 2, acquisition time 2 min 6 s (21). A 3D multi-echo Dixon protocol was employed to acquire fat-only and water-only images to calculate the fat fraction: VIBE sequence; 6 echoes; voxel size 1.1 × 1.1 × 3.0 mm^3; reconstructed matrix = 320 × 190 × 96; TR = 20 ms; TE = 1.41, 2.87, 4.33, 5.79, 7.25, and 8.71 ms; flip angle = 12°; acceleration factor 2; and acquisition time 4 min 49 s (22). DTI with a 2D echo-planar imaging (EPI) sequence was used to acquire maps of the ADC and the FA: 46 consecutive sagittal slices, voxel size 2.5 × 2.5 × 2.5 mm^3, reconstructed matrix = 128 × 52 × 46, TR = 4.0 s, TE = 49 ms, monopolar diffusion gradients in 30 directions with b-values of b = 0 and 700 s/m^2, fat-saturated, partial Fourier acquisition factor 7/8, acquisition time 2 min 18 s. Maps of T_2, ADC, FA, and fat fraction (defined as the signal intensity of the fat-only images divided by the sum of the signal intensities of fat-only and water-only images) were calculated online by the scanner software.

Image analysis was performed in the BTX-treated legs only. For each patient, segmentation of hyperintense regions of interest (ROIs) on T_2w images (TE = 34 ms, from the multi-spin-echo acquisition) of the second MRI time point (6w post-BTX) was performed by an experienced medical imaging technologist (PM) with ITK SNAP 3.6 using thresholding, region growing, and manual correction. A reference ROI (volume 1–2 cm^3) was drawn manually in the anterior muscle compartment (not BTX treated) with ITK snap. MR images were co-registered with FSL 5.0.11. First, within the same time point, the T_2w image (TE = 34 ms) and the Dixon acquisition image (TE = 2.87 ms) were co-registered on the b0 image (=image without diffusion weighting) of the DTI dataset. Second, between the time points, T_2w images pre-BTX and 12w post-BTX were co-registered on the T_2w image 6w post-BTX. The calculated transformation matrices were used to transfer both the hyperintense and reference ROIs onto the T_2 maps, ADC maps, FA maps, and fat fraction maps of each time point. The ROI mean values were calculated for each map.

The whole muscle delineation was done on the Dixon acquisition images (TE = 2.87 ms) at 6w time point by an experienced radiologist (TAD) using an in-house developed segmentation tool. Then the volumes of the GM, GL, and S in the treated legs were calculated. For calculating the relative ROI volume, the volume of the hyperintense ROI (after transfer on the Dixon acquisition images) was divided by the summed-up volume of the target muscles (GM+GL+S for patients 1 to 5; GM+GL for patient 6).

RESULTS

Clinical Examination

Clinical examinations did not show any relevant changes after treatment in five of six patients: there were no changes larger than 1 in the scores for spasticity and muscle strength (MMT) and no changes over 10° for RoM between pre-injection and 6w or 12w post-injection (**Table 1**). Only patient 5 showed an improvement in two out of six parameters (passive RoM for dorsiflexion and spasticity for the plantarflexor muscles).

Magnetic Resonance Imaging

Hyperintensity on T_2w scans and considerably increased T_2 (by approximately 10 ms compared with baseline) were observed in all patients in the GM and—more pronounced—in the S at the injection sites at 6w and 12w post-injection (**Figures 1**, **2**). T_2 in the reference region was constant over time, with a group mean value across six patients (±standard deviation STD across six patients) of 35 ± 1, 36 ± 2, and 36 ± 2 ms at time points pre-BTX, 6w, and 12w, respectively (**Figure 3**, top row, and **Figure 4**). In the hyperintense ROI, the mean value (±STD) across six patients increased from 35 ± 1 ms pre-BTX to 45 ± 2 ms at 6w and 44 ± 2 ms at 12w post-BTX injection. There was no hyperintensity in

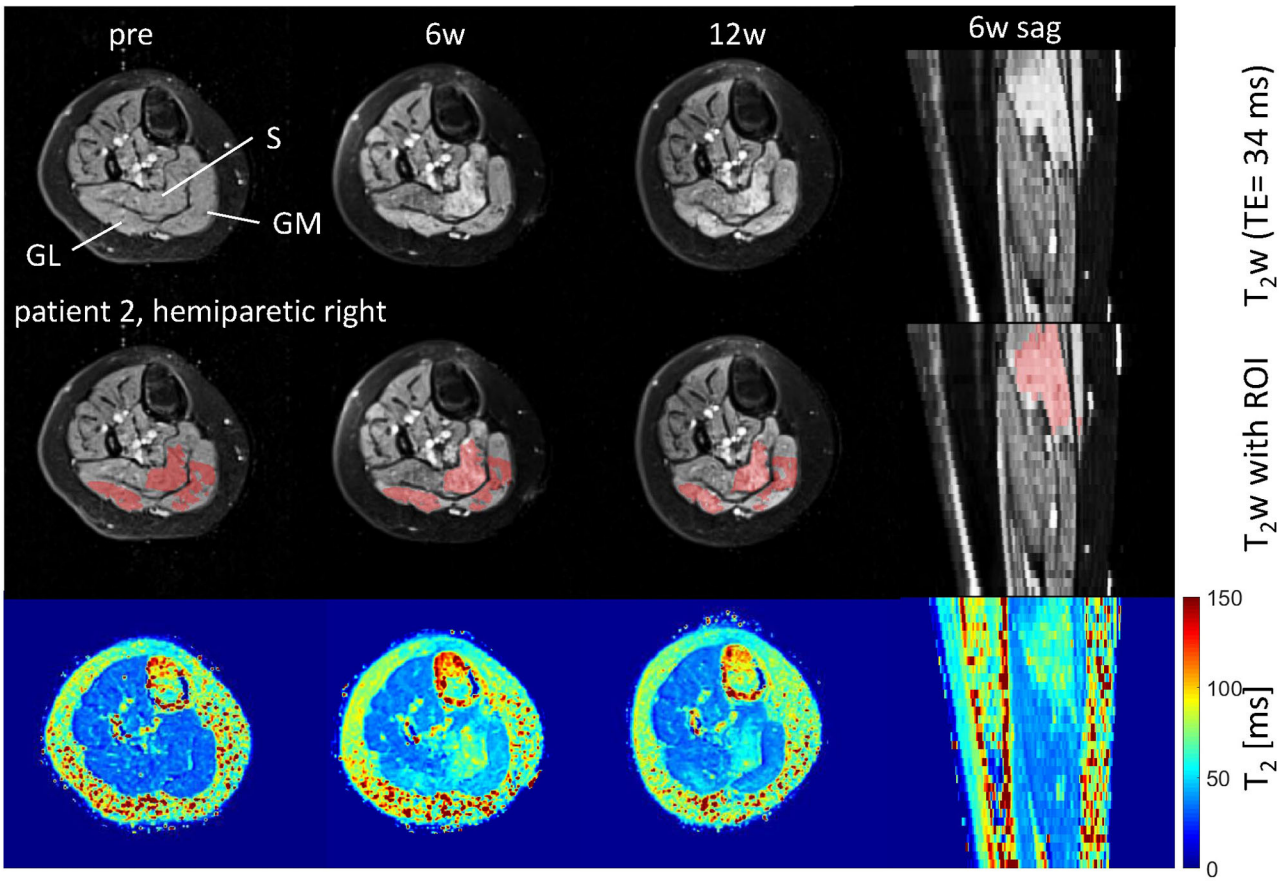


FIGURE 1 | Axial and sagittal T_2-weighted images in the calf of patient 2 (hemiparetic) pre-BTX and 6 weeks (6w) and 12 weeks (12w) post-BTX (top row) showing hyperintensity at the sites of injection in the soleus (S), gastrocnemius medialis, and lateralis (GM, GL). A region of interest (ROI) comprising the hyperintense regions was segmented and is shown as red overlay in the middle row. The increased T_2 post-BTX can be seen in the calculated T_2 maps (bottom row). BTX, botulinum toxin A; y, years; T_2w, T_2-weighted; TE, echo time; sag, sagittal.

the S of patient 6, as expected, since this patient was not treated in the S (**Figure 2**, right column).

The T_2 effect, however, was spatially limited and was only observed in parts of the muscles. That was mainly obvious in the S when looking at the extent of the hyperintense ROI (**Figure 1**, middle row). The hyperintense ROI covered 9–15% of the total GM, GL, and S volumes in all patients except in patient 6, and 43% of the volume of the GM+GL in patient 6 treated the GM and GL only (**Table 2**).

ADC did not change in these hyperintense ROIs (**Figures 3**, **4**): group mean values ± STD were 1.37 ± 0.05 * 10^{-3} mm^2/s pre-BTX, 1.43 ± 0.09 * 10^{-3} mm^2/s at 6w, and 1.41 ± 0.08 * 10^{-3} mm^2/s at 12w. Also in the reference ROIs, there was no change in the mean ADC: 1.46 ± 0.09 * 10^{-3} mm^2/s pre-BTX, 1.44 ± 0.08 * 10^{-3} mm^2/s at 6w, and 1.47 ± 0.04 * 10^{-3} mm^2/s at 12w. FA group mean ± STD was 0.30 ± 0.03 pre-BTX and increased to 0.34 ± 0.02 at 6w and 0.36 ± 0.03 at 12w. The FA values in the reference ROIs were stable: 0.26 ± 0.02 pre-BTX, 0.26 ± 0.02 at 6w, and 0.27 ± 0.03 at 12w.

There was a trend of an increase in the fat fraction in the hyperintense ROI at 12w: group mean values ± STD were 9.1 ± 3.7% pre-BTX, 10.6 ± 4.4% at 6w, and 13.7 ± 5.4% at 12w. Patient 2 was overweight (BMI > 85^{th} percentile) and showed higher fat fraction values than the rest of the group, also in the reference ROIs (top data points in **Figure 4**, top right).

DISCUSSION

The effect of BTX injections into the calf muscles of pediatric CP patients was quantitatively assessed using MRI over 12 weeks. Our study showed substantial elevation of T_2 values at the BTX injection sites in pediatric CP patients up to 12w post-treatment. It is not a direct visualization of BTX bolus, as reported by Elwischger et al. (23). In their study in healthy and spastic biceps brachii muscles in adults, the convection of the fluid depot was visualized as hyperintense areas in images at the injection site in a few minutes after BTX injection. The bolus expanded primarily along the muscle fibers and dispersed to a longer distance in healthy compared with spastic muscles. The distribution process

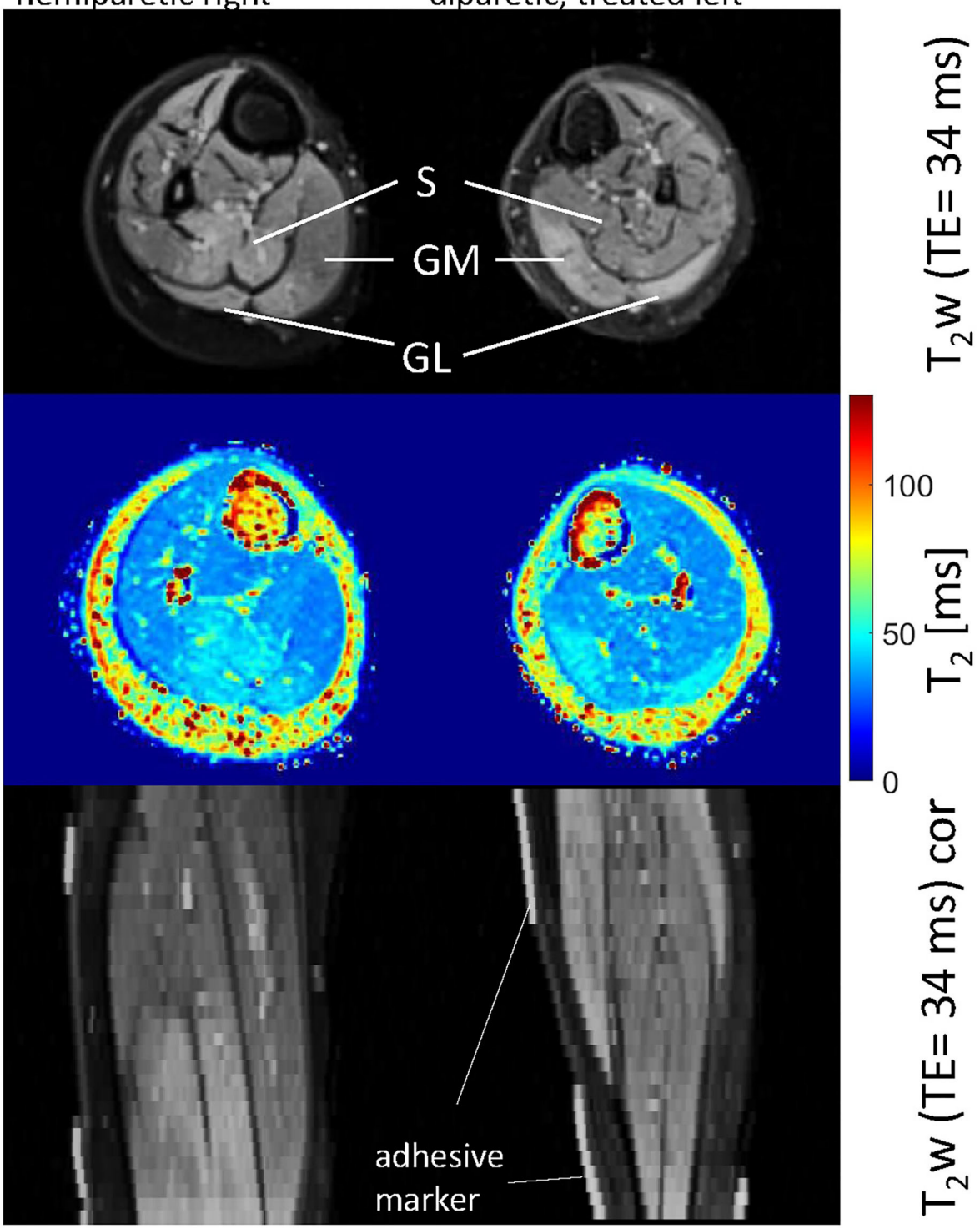


FIGURE 2 | Axial and coronal T_2-weighted images in the calves of patient 3 (hemiparetic) and patient 6 (diparetic) 6 weeks post-BTX (top and bottom row) showing hyperintensity at the sites of injection in the soleus (S), gastrocnemius medialis, and lateralis (GM, GL) for patient 3, and the GM and GL for patient 6. The increased T_2 post-BTX can be seen in the calculated T_2 maps (middle row). BTX, botulinum toxin A; y, years; T_2w, T_2-weighted; TE, echo time; cor, coronal.

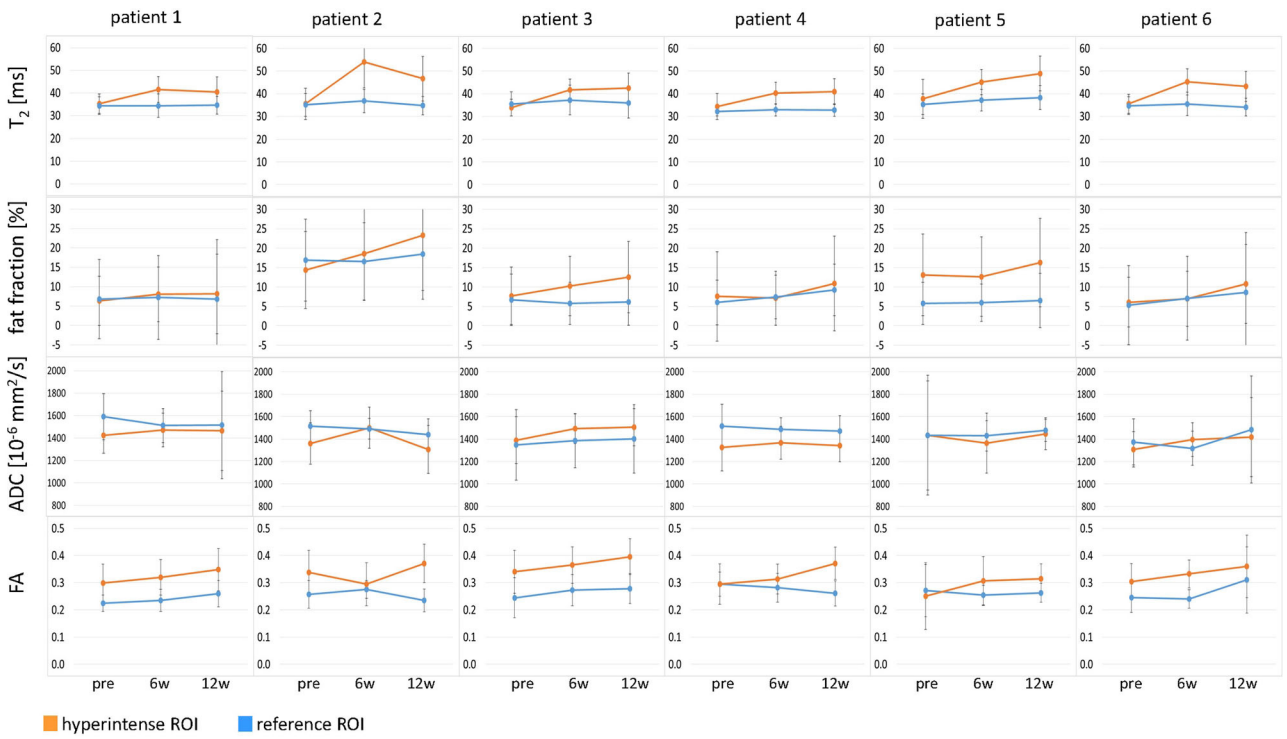


FIGURE 3 | Time courses of T_2, fat fraction, apparent diffusion coefficient (ADC), and fractional anisotropy (FA) in all six patients at time points pre-BTX, 6 weeks (6w), and 12 weeks (12w) post-BTX. The values are displayed as mean values in the hyperintense ROI (orange) and the reference ROI (blue). The error bars show the standard deviation in the ROIs. ROI, region of interest; BTX, botulinum toxin A.

was not completed at the second MRI at 11 min after injection, and therefore, it would be of interest to know the effect of BTX on muscle tissue after a more extended period. In our study, the bolus had already dispersed by the time of the MRI exam at 6w after the BTX injections. Our results are in line with O'Dell et al. (24). In their study, 200 units of BTX were injected in various leg muscles of adult patients who suffered a stroke (total dose of 200 units). An increase in T_2 was observed 2 and 3 months after BTX injection. Schroeder et al. (25) observed a hyperintensity in T_2w images up to 1 year in the gastrocnemius muscles of two healthy adults injected with a single dose of BTX (75 units). As already pointed out in these two studies, the increase in T_2 is presumably caused by an increase in the extracellular space (ECS) around the atrophic muscle fibers after treatment. In Schroeder et al.'s study, the treated muscles' biopsies showed increased connective tissue and enlargement in ECS (25).

In the absence of fat-suppression techniques, global T_2 increases with an increase in fat content, with fat having a longer T_2 than water. In our study, we acquired the T_2 maps with fat suppression. However, residual fat signal from an imperfect fat suppression may influence the measured T_2 values (26). Since the fat fraction did not change at 6w post-BTX, the elevated T_2 represents a T_2 increase in muscle water and is presumably caused by an ECS increase.

Regarding the diffusion results, our ADC values of approximately $1.4 * 10^{-3}$ mm^2/s are in the same range as in adult calf muscles (11). The FA values around 0.3 to 0.4 are in line with previously published results in healthy adult volunteers (27). We could not detect an increase in ADC as one would expect with an increased ECS. However, changes in muscle fiber size (i.e., fiber diameter) also have to be taken into account. BTX injections cause a decrease in muscle fiber size and, therefore, atrophy of the muscle (28). In the work by Berry et al. (29), a simulation of a standard single-echo DTI experiment in BTX-treated muscle with an increased ECS and decreased fiber size resulted in minimal changes in ADC, as ADC decreases with fiber size. As shown in the same study (29), FA is more sensitive to fiber size changes than ADC, as it increases with decreasing fiber size. The upper simulation of the BTX-treated muscle with reduced fiber size and increased ECS resulted in an increased FA. In our study, we observed a slight increase in FA over time post-BTX. Our results are also in line with a study in a surgical denervation model in rats causing muscle atrophy: no significant difference in muscle ADC was found between the control group and the denervated group, while FA increased in the atrophic muscle (30). These studies and our observations indicate that an increase in ECS and a simultaneous decrease of muscle fiber size by the application of BTX may counteract the effect on ADC, thus resulting in stable ADC values, whereas FA increases.

Concerning the fat quantification, we only observed a trend for an increase in fat fraction at 12w post-BTX. The fat fraction values had a large standard deviation across the ROIs. Chemical

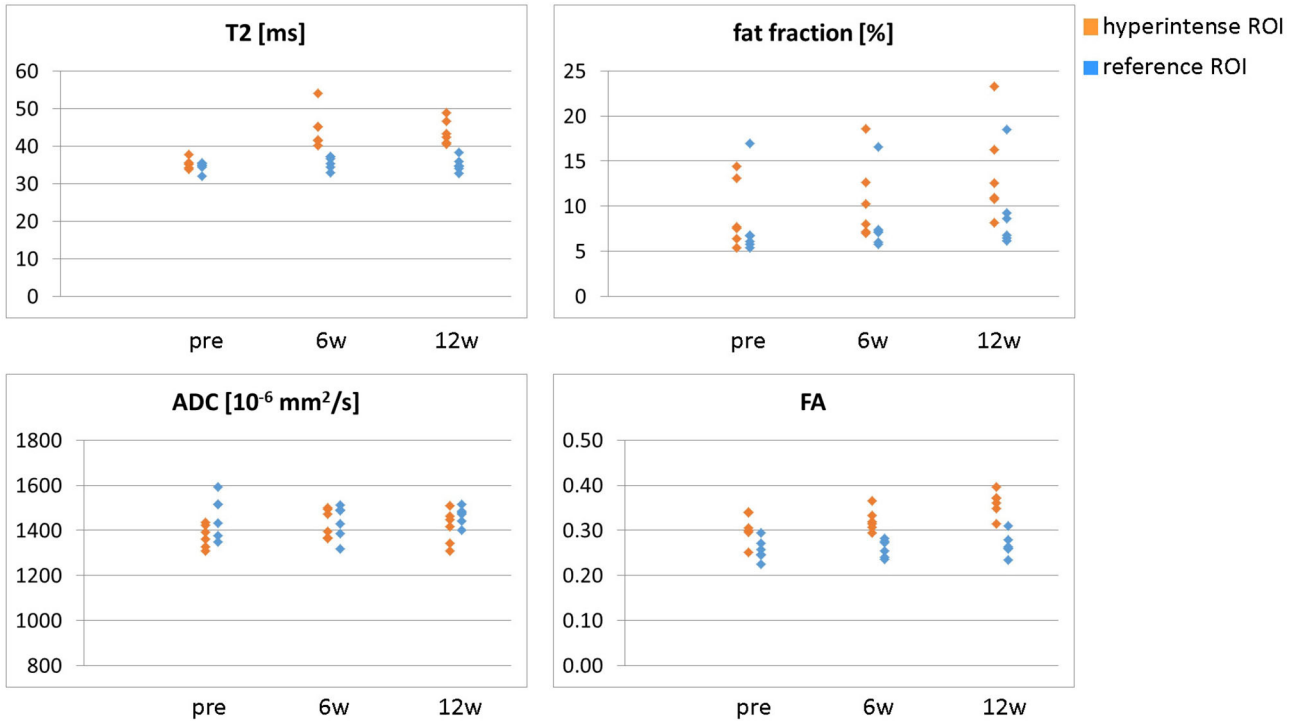


FIGURE 4 | Overview of all ROI mean values (T_2, fat fraction, ADC, and FA) in the hyperintense (orange) and reference ROIs (blue) for all six patients at time points pre-BTX, 6 weeks (6w), and 12 weeks (12w) post-BTX. ROI, region of interest; BTX, botulinum toxin A; ADC, apparent diffusion coefficient; FA, fractional anisotropy.

TABLE 2 | Volume of treated muscles and percentage of tissue with elevated T_2, i.e., relative ROI size at 6 weeks post-BTX.

Patient number	Treated muscles	Total volume of treated muscles (cm^3)	ROI size (cm^3)	Percentage of tissue with elevated T_2 = relative ROI size %
1	GM, GL, S	225	19.6	9
2	GM, GL, S	223	32.2	14
3	GM, GL, S	231	32.7	14
4	GM, GL, S	166	24.3	15
5	GM, GL, S	203	21.3	10
6	GM, GL	58	25.2	43

ROI, region of interest; GM, m. gastrocnemius medialis; GL, m. gastrocnemius lateralis; S, m. soleus; BTX, botulinum toxin A.

denervation by BTX—like denervation in general—leads to a partial replacement of muscle fibers by fat (3). However, the increase in fat content is only small after BTX treatment. A mild increase in the number of fat cells 12 months after a single dose of BTX was reported in biopsy calf muscle specimens (25). A small amount of lipids in the range of 1% relative to total muscle cross-sectional area (compared with almost 0% in controls) was detected in sections of rat muscle 3 months after a single dose of BTX (28). Such a slight change in fat fraction was challenging to detect in our study. Possibly, the fat fraction may further increase and be better detectable after 12w post-BTX until the BTX effect wears off and recovery of muscle morphology sets in Multani et al. (3).

As expected in our patient cohort, clinical examination showed no significant functional improvement after BTX. In this case, the intention of the injection was not the treatment of the equinus but to test the weakening of the muscle for a risk evaluation of the planned lengthening surgery (13). Therefore, their toe-walking remained unchanged after injection. This is in line with the general observation that a response to equinus treatment with BTX is barely detectable in children over 6 years (3). The patients in this study all showed fixed calf muscle contractures in the clinical examination. BTX thus would only have an effect on the additional functional part of the equinus and on stability. Muscle strength as assessed with MMT did not change in our patient cohort after injection. The decision to proceed with the planned lengthening surgery was based on the post-BTX gait analysis. Since the gait pattern did not deteriorate after BTX injection, all six patients were scheduled for surgery, as poor surgery outcomes due to further muscle weakening were considered unlikely.

T_2w images and T_2 maps showed hyperintensity and increased quantitative values, respectively, but only in parts and not in the muscles' complete cross section. Based on animal and human data and clinical experience, BTX diffusion is likely to occur up to a distance of 4.5–5 cm from the injection site (to some extent, even into adjacent muscles). This depends on the dosage

and volume injected (31) and on the muscle fiber structure and presence of spasticity (23). The hyperintense region's extent in our study in CP patients with muscle contractures was smaller than this maximum distance. The same review (31) also points out that directly detecting BTX spread in the human muscle is difficult. T_2-weighted imaging may be a potential tool to assess this spread non-invasively.

Our study has some limitations. The number of patients was small. Therefore, we refrained from using statistical testing for differences in the MRI results and solely used descriptive statistics in our preliminary report. Due to the relatively large standard deviations in the ROIs, a clear interpretation of the fat fraction and diffusion maps was difficult. However, the trends in the time courses after treatment were compelling and in line with the existing literature (e.g., simulations, biopsy data, and animal studies). The hyperintensity on T_2w images and T_2 maps up to 12 weeks after BTX treatment was clearly visible and confirmed the findings of the two published pilot studies in adults (24, 25). We performed this pilot study in children over 8 years of age since they could follow instructions in the scanner and lie still without sedation. It would be interesting to perform the experiments in younger children without fixed contractures, who are the main target group for BTX treatment, but then sedation might be necessary to avoid motion artifacts in the MR images. We only tested one specific treatment protocol with dose, dilution, and position of injection sites suited for the preoperative testing but not optimized for equinus treatment. The analysis of more patients within the current study and more studies on this topic are necessary to draw firmer conclusions.

The method may be applied in future studies with other BTX treatment protocols to quantify BTX-induced effects on muscle tissue and to investigate the spatial BTX distribution. A spatially limited extent of the BTX-induced effect, i.e., if the spread of BTX does not fully cover the zone of the motor end plates in the muscle (31), may help to explain a limited treatment effect on the gait and on the clinical scores of CP patients. Based on the MRI results, a revision of the applied treatment protocol regarding dose, dilution, or injections sites may be considered. The recommended total dose was reduced in the last years after serious side effects, especially in severely affected CP patients (2, 15). As discussed in a recent review (31), in relatively large muscles like leg muscles, it may be beneficial to use higher doses or higher dilutions (if a dose increase is not beneficial) and divide the dose over multiple injection sites.

In conclusion, our study shows that MRI is a useful tool for monitoring the spatial distribution of BTX treatment effects. It is especially well-suited to non-invasively determine BTX-induced changes and therapeutic responses in a pediatric population. To our knowledge, our study is the first to employ MRI to monitor BTX treatment effects in children. Using the proposed method as a monitoring tool in future studies will provide more detailed insight into the therapeutic effect of BTX in muscles of CP patients and may influence the treatment process (e.g., injection protocol).

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because restrictions apply to DICOM data originating from the MRI scanner. Requests to access the datasets should be directed to CW, claudia.weidensteiner@unibas.ch.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission Nordwest- und Zentralschweiz, Basel, Switzerland. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

RB, ER, CW, OB, and MG designed the study. RB supervised the project. CW, PM, and TH performed MRI experiments. KB-S, ED, and JR performed and analyzed clinical examinations. CW, TA, and PM analyzed MRI data. XD, FS, and PM developed analysis tools. XD and FS assisted with MRI experiments. CW wrote the manuscript. RB, OB, XD, FS, ED, JR, PM, TA, and MG revised and edited the manuscript. The final manuscript version was approved by all authors.

FUNDING

This study was supported by the Swiss National Science Foundation (Grant No. 173292).

REFERENCES

1. Graham HK, Rosenbaum P, Paneth N, Dan B, Lin J-P, Damiano DL, et al. Cerebral palsy. *Nat Rev Dis Primer*. (2016) 2:15082. doi: 10.1038/nrdp.2016.5
2. Strobl W, Theologis T, Brunner R, Kocer S, Viehweger E, Pascual-Pascual I, et al. Best clinical practice in botulinum toxin treatment for children with cerebral palsy. *Toxins*. (2015) 7:1629–1648. doi: 10.3390/toxins7051629
3. Multani I, Manji J, Hastings-Ison T, Khot A, Graham K. Botulinum toxin in the management of children with cerebral palsy. *Paediatr Drugs*. (2019) 21:261–81. doi: 10.1007/s40272-019-00344-8
4. Van Campenhout A, Molenaers G. Localization of the motor endplate zone in human skeletal muscles of the lower limb: anatomical guidelines for injection with botulinum toxin. *Dev Med Child Neurol*. (2011) 53:108–19. doi: 10.1111/j.1469-8749.2010.03816.x
5. Maillard SM, Jones R, Owens C, Pilkington C, Woo P, Wedderburn LR, Murray KJ. Quantitative assessment of MRI T2 relaxation time of thigh muscles in juvenile dermatomyositis. *Rheumatol Oxf Engl*. (2004) 43:603–8. doi: 10.1093/rheumatology/keh130
6. Marty B, Baudin P-Y, Reyngoudt H, Azzabou N, Araujo ECA, Carlier PG, de Sousa PL. Simultaneous muscle water T2 and fat fraction mapping using transverse relaxometry with stimulated echo compensation. *NMR Biomed*. (2016) 29:431–43. doi: 10.1002/nbm.3459
7. Arpan I, Forbes SC, Lott DJ, Senesac CR, Daniels MJ, Triplett WT, Deol JK, et al. T2 mapping provides multiple approaches to characterize muscle involvement in neuromuscular diseases: a cross-sectional study of lower leg

muscles in 5–15 year old boys with Duchenne Muscular Dystrophy. *NMR Biomed*. (2013) 26:320–8. doi: 10.1002/nbm.2851
8. Bonati U, Hafner P, Schädelin S, Schmid M, Naduvilekoot Devasia A, Schroeder J, et al. Quantitative muscle MRI: A powerful surrogate outcome measure in Duchenne muscular dystrophy. *Neuromuscul Disord NMD*. (2015) 25:679–85. doi: 10.1016/j.nmd.2015.05.006
9. Stouge A, Khan KS, Kristensen AG, Tankisi H, Schlaffke L, Froeling M, et al. MRI of skeletal muscles in participants with type 2 diabetes with or without diabetic polyneuropathy. *Radiology*. (2020) 297:608–19. doi: 10.1148/radiol.2020192647
10. Oudeman J, Nederveen AJ, Strijkers GJ, Maas M, Luijten PR, Froeling M. Techniques and applications of skeletal muscle diffusion tensor imaging: a review. *J Magn Reson Imaging JMRI*. (2016) 43:773–88. doi: 10.1002/jmri.25016
11. Zaraiskaya T, Kumbhare D, Noseworthy MD. Diffusion tensor imaging in evaluation of human skeletal muscle injury. *J Magn Reson Imaging*. (2006) 24:402–8. doi: 10.1002/jmri.20651
12. Hooijmans MT, Damon BM, Froeling M, Versluis MJ, Burakiewicz J, Verschuuren JJGM, et al. Evaluation of skeletal muscle DTI in patients with duchenne muscular dystrophy. *NMR Biomed*. (2015) 28:1589–97. doi: 10.1002/nbm.3427
13. Rutz E, Hofmann E, Brunner R. Preoperative botulinum toxin test injections before muscle lengthening in cerebral palsy. *J Orthop Sci Off J Jpn Orthop Assoc*. (2010) 15:647–53. doi: 10.1007/s00776-010-1509-6
14. Palisano RJ, Rosenbaum P, Bartlett D, Livingston MH. Content validity of the expanded and revised gross motor function classification system. *Dev Med Child Neurol*. (2008) 50:744–50. doi: 10.1111/j.1469-8749.2008.03089.x
15. Heinen F, Desloovere K, Schroeder AS, Berweck S, Borggraefe I, van Campenhout A, et al. The updated European consensus 2009 on the use of Botulinum toxin for children with cerebral palsy. *Eur J Paediatr Neurol EJPN Off J Eur Paediatr Neurol Soc*. (2010) 14:45–66. doi: 10.1016/j.ejpn.2009.09.005
16. Hamjian JA, Walker FO. Serial neurophysiological studies of intramuscular botulinum-A toxin in humans. *Muscle Nerve*. (1994) 17:1385–92. doi: 10.1002/mus.880171207
17. Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V, et al. Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschr Kinderheilkd*. (2001) 149:807–18. doi: 10.1007/s001120170107
18. Hefti F. *Pediatric Orthopedics in Practice*. Berlin; Heidelberg: Springer-Verlag (2007).
19. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Phys Ther*. (1987) 67:206–7. doi: 10.1093/ptj/67.2.206
20. Hislop HJ, Montgomery J. *Daniels and Worthingham's Muscle Testing: Techniques of Manual Examination*. St. Louis, MO.: Saunders Elsevier (2007).
21. Hilbert T, Sumpf TJ, Weiland E, Frahm J, Thiran J-P, Meuli R, et al. Accelerated T2 mapping combining parallel MRI and model-based reconstruction: GRAPPATINI. *J Magn Reson Imaging JMRI*. (2018) 48:359–68. doi: 10.1002/jmri.25972
22. Henninger B, Zoller H, Kannengiesser S, Zhong X, Jaschke W, Kremser C. 3D Multiecho dixon for the evaluation of hepatic iron and fat in a clinical setting. *J Magn Reson Imaging*. (2017) 46:793–800. doi: 10.1002/jmri.25630
23. Elwischger K, Kasprian G, Weber M, Meyerspeer M, Linder C, Auff E, et al. Intramuscular distribution of botulinum toxin–visualized by MRI. *J Neurol Sci*. (2014) 344:76–9. doi: 10.1016/j.jns.2014.06.028
24. O'Dell MW, Villanueva M, Creelman C, Telhan G, Nestor J, Hentel KD, et al. Detection of botulinum toxin muscle effect in humans using magnetic resonance imaging: a qualitative case series. *PMR*. (2017) 9:1225–35. doi: 10.1016/j.pmrj.2017.04.014
25. Schroeder AS, Ertl-Wagner B, Britsch S, Schröder JM, Nikolin S, Weis J, et al. Muscle biopsy substantiates long-term MRI alterations one year after a single dose of botulinum toxin injected into the lateral gastrocnemius muscle of healthy volunteers. *Mov Disord Off J Mov Disord Soc*. (2009) 24:1494–503. doi: 10.1002/mds.22661
26. Carlier PG, Marty B, Scheidegger O, Loureiro de Sousa P, Baudin P-Y, Snezhko E, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging and spectroscopy as an outcome measure for clinical trials. *J Neuromuscul Dis*. (2016) 3:1–28. doi: 10.3233/JND-160145
27. Galbán CJ, Maderwald S, Uffmann K, de Greiff A, Ladd ME. Diffusive sensitivity to muscle architecture: a magnetic resonance diffusion tensor imaging study of the human calf. *Eur J Appl Physiol*. (2004) 93:253–62. doi: 10.1007/s00421-004-1186-2
28. Minamoto VB, Suzuki KP, Bremner SN, Lieber RL, Ward SR. Dramatic changes in muscle contractile and structural properties after two botulinum toxin injections. *Muscle Nerve*. (2015) 52:649–57. doi: 10.1002/mus.24576
29. Berry DB, Regner B, Galinsky V, Ward SR, Frank LR. The relationships between tissue microstructure and the diffusion tensor in simulated skeletal muscle. *Magn Reson Med*. (2018) 80:317–29. doi: 10.1002/mrm.26993
30. Saotome T, Sekino M, Eto F, Ueno S. Evaluation of diffusional anisotropy and microscopic structure in skeletal muscles using magnetic resonance. *Magn Reson Imaging*. (2006) 24:19–25. doi: 10.1016/j.mri.2005.09.009
31. Sätilä H. Over 25 years of pediatric botulinum toxin treatments: what have we learned from injection techniques, doses, dilutions, and recovery of repeated injections? *Toxins*. (2020) 12:440. doi: 10.3390/toxins12070440

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 21 May 2021
doi: 10.3389/fneur.2021.613483

Patients With Becker Muscular Dystrophy Have Severe Paraspinal Muscle Involvement

Aisha M. Sheikh, Karen Rudolf, Josefine de Stricker Borch, Tahmina Khawajazada, Nanna Witting and John Vissing*

Copenhagen Neuromuscular Center, Department of Neurology, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

Introduction: Paraspinal muscles are important for gross motor functions. Impairment of these muscles can lead to poor postural control and ambulation difficulty. Little knowledge exists about the involvement of paraspinal muscles in Becker muscular dystrophy.

Objective: In this cross-sectional study, we investigated the involvement of paraspinal muscles with quantitative trunk strength measure and quantitative muscle MRI.

Methods and Materials: Eighteen patients with Becker muscular dystrophy underwent trunk, hip, and thigh strength assessment using a Biodex dynamometer and an MRI Dixon scan. Fourteen age- and body mass index-matched healthy men were included for comparison.

Results: Muscle fat fraction (FF) of the paraspinal muscles (multifidus and erector spinae) was higher in participants with Becker muscular dystrophy vs. healthy controls at all three examined spinal levels (C6, Th12, and L4/L5) ($p < 0.05$). There was a strong and inverse correlation between paraspinal muscle FF and trunk extension strength ($\rho = -0.829$, $p < 0.001$), gluteus maximus FF and hip extension strength ($\rho = -0.701$, $p = 0.005$), FF of the knee extensor muscles (quadriceps and sartorius) and knee extension strength ($\rho = -0.842$, $p < 0.001$), and FF of the knee flexor muscles (hamstring muscles) and knee flexion strength ($\rho = -0.864$, $p < 0.001$). Fat fraction of the paraspinal muscles also correlated with muscle FF of the thigh muscles and lower leg muscles.

Conclusion: In conclusion, patients with Becker muscular dystrophy demonstrate severe paraspinal muscular involvement indicated by low back extension strength and high levels of fat replacement, which parallel involvement of lower limb muscles. Assessment of paraspinal muscle strength and fat replacement may serve as a possible biomarker for both the clinical management and further study of the disease.

Keywords: paraspinal muscles, Becker muscular dystrophy, quantitative muscle MRI, quantitative trunk strength, fat fraction

OPEN ACCESS

Edited by:
Francesco Santini, University of Basel, Switzerland

Reviewed by:
Yun Yuan, Peking University First Hospital, China
Melissa Hooijmans, VU University Medical Center, Netherlands

****Correspondence:***
Aisha M. Sheikh
aisha.munawar.sheikh@regionh.dk

Specialty section:
This article was submitted to Applied Neuroimaging, a section of the journal Frontiers in Neurology

Received: *02 October 2020*
Accepted: *27 April 2021*
Published: *21 May 2021*

Citation:
Sheikh AM, Rudolf K, de Stricker Borch J, Khawajazada T, Witting N and Vissing J (2021) Patients With Becker Muscular Dystrophy Have Severe Paraspinal Muscle Involvement. Front. Neurol. 12:613483. doi: 10.3389/fneur.2021.613483

INTRODUCTION

Paraspinal muscles consist of erector spinae, which has a superficial segment consisting of the iliocostalis, spinalis, and longissimus muscles, and a deep segment consisting of the multifidus muscle (1). Paraspinal muscles provide trunk stability and mobility and are vital in motor tasks such as rising from a chair and walking. Impairment of these muscles can impact activities of daily living and lead to poor posture, pain, and endurance.

Clinical assessment of paraspinal muscles is challenging because muscle volume is problematic to assess visually and testing the strength of these muscles is inherently difficult. Consequently, the involvement of paraspinal muscles is largely unexplored in patients with a variety of neuromuscular diseases (2). Currently, the prone Biering-Sørensen test is widely used to assess isometric back extension strength. This test uses a hand-held dynamometer which is held over the interscapulum region of the back by the examiner, against which the patient is asked to perform a maximal counter pressure. The test is dependent on the degree of resistance on the hand-held dynamometer generated by the examiner and has been shown to have poor reliability (3) and a ceiling effect as many test individuals overcome the strength of the examiner. An adaptation to the widely used Biodex system to assess muscle strength has introduced a new isometric trunk dynamometer. This system differs from other strength measure tests in that strength is measured from a seated position and is not examiner-dependent. A seated position allows for a feasible measurement of strength, allowing the examiner to assess the strength of weaker patients and removing the challenges of individuals frequently outperforming the examiner's strength.

Magnetic resonance imaging (MRI), using a Dixon technique to assess muscle fat fractions, provides a quantitative method to validate disease distribution and severity of muscle involvement (4–6) and has been shown to correlate with functional assessments and disease progression in several muscular dystrophies (7–12). Previous MRI studies have demonstrated that patients with Becker muscular dystrophy (BMD) exhibit high levels of fat replacement, particularly the hamstrings, adductors, quadriceps, and gastrocnemius muscles (7, 8, 13).

In patients with BMD, a systematic investigation of the paraspinal muscles has not been performed yet. Using a combination of strength measures and quantitative MRI (qMRI), we investigated paraspinal muscle involvement in patients with BMD.

MATERIALS AND METHODS

Study Design and Participants

This cross-sectional study was conducted at the Copenhagen Neuromuscular Center at the National University Hospital, Rigshospitalet in Copenhagen, Denmark from April 2018 to June 2020 in accordance with the declaration of Helsinki and was approved by the Danish National Committee on Health Research Ethics (approval number: H-16030358).

Inclusion criteria were (1) 18 years of age or older, because Copenhagen Neuromuscular Center is a center for adult patients, (2) genetically verified BMD, (3) able to comprehend and adhere to participation requirements, and (4) ability to stand with or without an assistive device or support on wall or furniture. Exclusion criteria were (1) conditions of the spine (i.e., severe scoliosis, past spine surgery) which could interfere with the results, and (2) contraindications to MRI.

Fifty-three patients with BMD from our center were potential candidates to participate in this study. Seven patients did not wish to participate for personal reasons, and it was not possible to obtain contact with nine patients. Another 19 patients were excluded due to exclusion criteria, 15 wheel-chair bound, one cognitively impaired, and three with contraindications to MRI. This left 18 patients who agreed to participate (**Table 1**). Fourteen age- and body mass index (BMI) -matched healthy men (HC) were included in the study for comparison of findings in patients with BMD. Quantitative muscle strength measure and qMRI were completed in a single visit. The participants were asked to refrain from non-habitual physical activity the day before the visit.

Before data collection, written informed consent was obtained from all participants. Written informed consent was also obtained from the participants for the publication of any potentially identifiable images or data included in this article.

Quantitative Muscle Strength Measure

The maximal voluntary isometric contraction was measured in the following order: (1) hip flexion and extension in supine position of the dominant leg, (2) knee extension and flexion in sitting position of the dominant leg, and (3) trunk extension and flexion in sitting position. Hip and thigh muscle strength was acquired using a stationary dynamometer (Biodex System 4 Pro, Biodex Medical Systems, Shirley, NY).

Trunk muscle strength was acquired using a Biodex Dual-position back Extension/Flexion attachment (model number 830-450). The anterior superior iliac spine was aligned with the attachment's fixed axis of rotation and back support was set at 100° hip angle. To minimize the influence of muscles from other parts of the body, the chest, and thighs were immobilized with Velcro straps, and the participants were asked to cross their arms in front of their chest during the test. Each participant was instructed to perform a maximal isometric trunk extension and trunk flexion. One contraction lasted 5 s. To ensure maximal contraction, two sub-maximal test trials were performed to familiarize the participants with the testing protocol for each position, followed by three trials of maximal contractions (interchangeably between extension and flexion) with 30 s of rest in between each contraction. Standardized verbal encouragement was provided to each participant during testing. To avoid muscle fatigue, each participant rested for 30 min between the lower limb muscle strength test and trunk muscle strength test.

MRI Data Acquisition

Images were acquired with a Siemens 3.0 Tesla Magnetom Verio scanner (Erlangen, Germany) at the Department of Radiology at the National University Hospital, Rigshospitalet in Copenhagen, Denmark. Each participant was examined in the headfirst supine

TABLE 1 | Demographics of the 18 participating patients with Becker muscular dystrophy.

Subject ID	Age	BMI	First symptom	Age at symptom onset	Symptoms at visit	Pathogenic variant	Disease duration
BMD 1	45	27.6	Difficulty running, stiffness	28	Severe difficulty walking and rising from a chair. Climbing stairs possible with support.	c.676_678del; p.(Lys226del)	17
BMD 2	50	27.5	Stiffness	41	Moderate to severe pain and joint stiffness.	Del26; p.(Val1145_Lys1201del)	9
BMD 3	36	22.6	Pain	10	Mild pain and joint stiffness.	Del45-48; p.(Glu2147_Gln2366del)	26
BMD 4	30	25.8	Fatigue	10	Moderate fatigue and muscle cramps.	c.6912+1G>T; p.(?)[§]	20
BMD 5	37	28.0	Unable to jump	2	Dependent on an electric wheelchair for most functions. Severe difficulty walking and rising from a chair. Climbing stairs not possible.	Del45-48; p.(Glu2147_Gln2366del)	35
BMD 6	32	25.1	Pain, muscle weakness	29	Severe difficulty walking, rising from a chair, and stair climbing. No use of assistive device.	c.1602 G>A; p.(?)[§]	3
BMD 7	27	23.8	Difficulty climbing stairs	24	Moderate difficulty walking, rising from a chair, and climbing stairs. No use of assistive device.	Del45-47; p.(Glu2147_Lys2304del)	3
BMD 8	33	23.6	Muscle cramps	5	Electric wheelchair for most functions. Able to stand for very short duration with support.	Del45-48; p.(Glu2147_Gln2366del)	28
BMD 9	18	29.9	Fatigue	1	Severe fatigue.	c.5632C>T; p.(Gln1878*)	17
BMD 10	38	23.6	Difficulty climbing stairs	22	Severe difficulty walking, rising from a chair, and climbing stairs. Occasionally uses cane for walking.	Del45-48; p.(Glu2147_Gln2366del)	16
BMD 11	38	25.6	Pain	6	Severe difficulty walking, rising from a chair, and climbing stairs. No use of assistive device	Del45-48; p.(Glu2147_Gln2366del)	32
BMD 12	29	27.4	Asymptomatic	N/A	N/A	Del45-47; p.(Glu2147_Lys2304del)	N/A
BMD 13	31	34.3	Muscle weakness, cramps	8	Severe difficulty walking, rising from a chair, and climbing stairs. No use of assistive device.	Del45-47; p.(Glu2147_Lys2304del)	23
BMD 14	25	24.8	Fatigue	6	Severe difficulty walking, rising from a chair, and climbing stairs. No use of assistive device.	Del45-49; p.(Glu2147_Lys2400del)	19
BMD 15	59	22.9	Difficulty climbing stairs	35	Severe difficulty walking, rising from a chair, and climbing stairs. No use of assistive device.	Del48; p.(Val2305_Gln2366del)	24
BMD 16	59	48.2	Pain, difficulty running	7	Electric wheelchair for most functions. Able to stand for very short duration with support.	Del45-47; p.(Glu2147_Lys2304del)	52
BMD 17	45	34.1	Pain, cramps	17	Pain and severe fatigue.	Del2; p.(Tyr11Phefs*7)	28
BMD 18	18	20.5	Difficulty with activities	4	Severe difficulty walking, rising from a chair, and climbing stairs. No use of assistive device.	Del44; p.(Arg2098Asnfs*16)	14

Table displays age, BMI, first symptom, age at symptom onset, age at diagnosis, symptoms at visit, pathogenic variant, and disease duration in years. BMI, body-mass-index.
[§]The variant c.1602G>A predicts a synonymous effect on the protein (p.(Lys534=)), but changes the last base in exon 13, and is predicted to disrupt the 5′ spice site for which the protein consequence cannot be predicted confidently (p.(?)).

position. The MRI protocol was composed of a three-plane localizer sequence followed by a 2-point Dixon sequence (Field of view (FOV) 400–450 mm; slice thickness 3.5 mm; distance factor 0%; repetition time (TR)/echo time (TE) 2.45 and 3.675/5.59 ms; flip angle of 9 deg.) and a T1-weighted sequence (FOV 400–450 mm; slice thickness 6.0 mm; distance factor 20%; TE/TR 19/650 ms; flip angle of 120°).

The 2-point Dixon sequence had an in-plane resolution of 2.1 × 4 × 3.5 and consisted of a multi-slice sequence consisting of 60 slices per sequence. Also, for the Dixon sequence, we used a low flip angle of 9 deg. to suppress T1 relaxation effects to avoid an overestimation of the fat contribution. The T1-weighted sequence was performed to obtain a visual representation of the muscles and was not included in the statistical analysis.

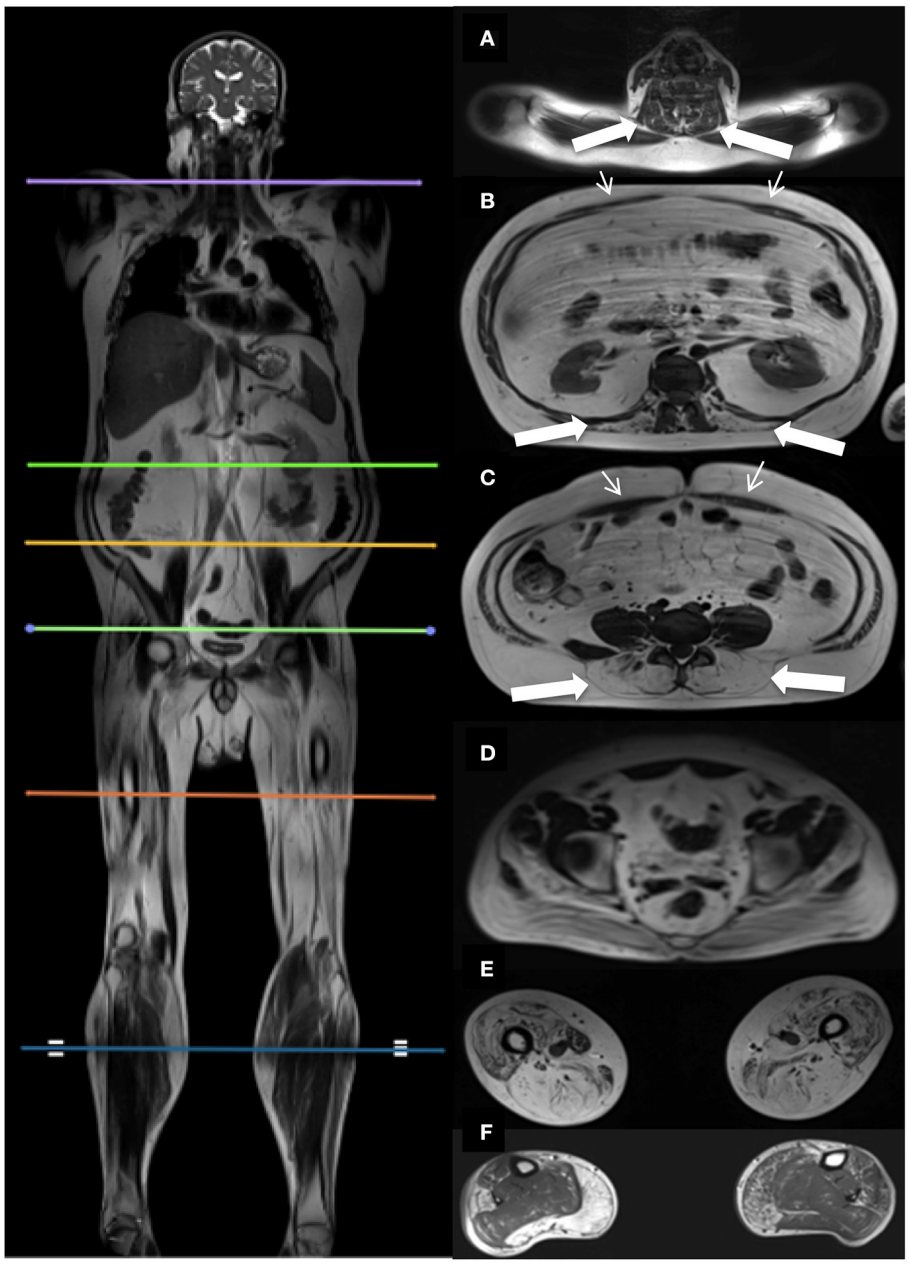


FIGURE 1 | Localizers for cross-sectional MR assessments, and corresponding cross-sectional images. Picture on the left shows the six cross-sectional slices at spinal level C6 **(A)**, Th12 **(B)**, L4/L5 **(C)**, S4 **(D)**, thighs **(E)**, and lower legs **(F)**. Images on the right show the corresponding cross-sectional images. Thick white arrows indicate the paraspinal muscles examined at spinal level C6 **(A)**, Th12 **(B)**, L4/L5 **(C)**. Thin white arrows indicate the abdominal muscles **(B,C)**.

Two body matrix coils and a peripheral leg coil were used for signal detection. The total scan time was approximately 40 min.

MRI Data Processing

Dixon sequences were used to quantify muscle fat fraction (FF) by defining a region of interest (ROI). Six cross-sectional slices were chosen for the investigation of muscle involvement (**Figure 1**). Horos software v. 3.3.6 was used to extract information from the 23 bilateral ROIs based on the raw Siemens data. The ROIs were drawn of the in-phase images and fat images and included the following muscles: Erector spinae at spinal level C6, Th12, and L4/L5, and multifidus at spinal level Th12 and L4/L5. Abdominal muscles were not mapped but instead inspected visually due to motion artifacts from respiration. Iliopsoas was mapped at spinal level L4/L5 and gluteus maximus at spinal level S4. Thigh muscles were mapped at mid-thigh, corresponding to 50% of the length of the femur, and included the following muscles: Rectus femoris (RF),

vastus lateralis (VL), vastus medialis (VM), vastus intermedius (VI), sartorius (SA), biceps femoris (BF), semitendinosus (ST), semimembranosus (SM), adductor muscles (AM), and gracilis (GR). At the widest section of the lower leg, corresponding to around 1/3 of tibia from the knee down, the following muscles were mapped: tibialis anterior (TA), peroneals (PER), tibialis posterior (TP), soleus (SOL), gastrocnemius lateralis (GL), and gastrocnemius medialis (GM).

We investigated the erector spinae muscle and multifidus muscle at three spinal levels to cover the important segments of the spine following previous MRI studies done in neuromuscular disorders (14–16). The thigh and lower leg muscles were also chosen based on earlier MRI studies (7, 8, 10–13, 17, 18).

Cross-sectional area (CSA) was determined from each ROI followed by a quantitative FF estimation. Mean muscle FF was expressed as

$$Percentage\ fat\ (0–100\%) = signal\ fat/(signal\ water + fat)$$ (8, 19)

Bilateral mean muscle FF was used in the analysis.

Contractile cross-sectional area (CCSA) was expressed as

$$CCSA = CSA \times (100\text{-}FF)/100$$ (20)

Muscle FF distribution was examined in the thigh and lower leg muscles by examining additional ROIs located 15 slices proximally and 15 slices distally from the center slice of the thigh and lower leg.

Statistical Analysis

Statistical analysis was performed using SPSS v22. The Mann-Whitney *U*-test was used to test the null hypothesis of no difference in paraspinal muscle FF between BMD and HC and the level of significance was set at $p \leq 0.05$.

The Friedman test was used to test for differences in FF distribution in the thigh and lower leg muscles. Wilcoxon test was performed with a Bonferroni adjustment on significant differences in FF distribution and the level of significance was set at $p \leq 0.017$. We report median values and interquartile ranges (Q1, Q3) unless otherwise stated.

Spearman correlation was used to evaluate the correlation between (1) trunk extension strength and paraspinal muscle FF in BMD, (2) hip strength and hip muscle FF in BMD, (3) thigh strength and thigh muscle FF in BMD, (4) paraspinal muscle FF and disease duration, where disease duration was defined as the number of years from symptom onset to the participant's age on the day of the visit, (5) paraspinal muscle FF and age in BMD and HC, (6) paraspinal muscle FF and muscle FF of the thigh muscles and lower leg muscles in BMD, and (7) contractile cross-sectional area and hip flexion strength of iliopsoas muscle, knee extension strength of knee extensor muscles (quadriceps and sartorius), and knee flexion strength of knee flexor muscles (hamstrings) in BMD and HC.

A global muscle FF was defined as the average of all ROIs of one cross-sectional slice and included the following: (1) Th12 (multifidus and erector spinae), (2) L4/L5 (multifidus and erector spinae, not including iliopsoas), (3) combined Th12 and L4/L5 (multifidus and erector spinae), (4) thigh muscles (RF, VL, VM, VI, SA, GR, BF, ST, SM, AM, GR), and (5) lower leg muscles (TA, PER, TP, SOL, GL, GM).

RESULTS

Becker muscular dystrophy participants had a mean age of 36.1 ± 11.9 years and HC participants were 36.4 ± 11.8 years. The mean BMI of BMD participants was 27.5 ± 6.3 and HC had a BMI of 26.6 ± 2.9.

Fat Fraction in BMD vs. HC

All examined paraspinal muscles exhibited severe muscle affection in BMD. When compared to HC, we found a statistically significant difference in median muscle FF on all three spinal levels (**Figure 2**). The median muscle FF at spinal level C6 erector spinae in participants with BMD was 18.4% (14.5, 21.9%) and in HC FF was 7.9% (6.8, 16.9%). At spinal level Th12, fat replacement of multifidus was 17.8% (15.8, 26.3%), and in erector spinae 62.2% (22.4, 71.1%), and in HC FF was 10.2% (7.7, 15.4%), and 9.7% (6.6, 15.9%), respectively. In the lumbar region, median FF was 41.9% (20.3, 72.7%) in multifidus and 50.9% (37.2, 78.9%) in erector spinae and 9.1% (5.8, 13.2%) and 18.8% (15.7, 26.9%) in HC, respectively. The median muscle FF of iliopsoas in BMD and HC was 17.7% (13.9, 27.3%) and 14.1% (9.2, 17.5%). Furthermore, a visual inspection of the abdominal muscles revealed that the muscles were well-preserved. In addition, MR images obtained on the asymptomatic participant showed modest intramuscular changes in erector spinae and multifidus at spinal level L4/L5.

We found no significant difference in muscle FF distribution of the hamstring muscles, tibialis anterior muscle, and peroneal muscles ($p > 0.05$). There was a significant difference in muscle FF distribution of the quadriceps muscle ($x^2 2 = 10.706$, $p = 0.005$) and calf muscle ($x^2 2 = 13.059$, $p = 0.001$). Median muscle FF of the three cross-sectional slices of the quadriceps were as follows: proximal slice 47.7% (15.8, 68%), center slice 52% (19.1, 70.2%), and distal slice 57.7% (22.4, 71.5%). There was a significant difference between the center slice and proximal slice ($Z = -2.485$, $p = 0.013$) and between the proximal slice and distal slice ($Z = -2.595$, $p = 0.003$), but there was no significant difference between the center slice and the distal slice ($Z = -0.923$, $p = 0.356$). Median muscle FF of the three cross-sectional slices of the calf muscle were as follows: proximal slice 18.2% (10.1, 43.3%), center slice 29.1% (14.9, 53.6%), and distal slice 27.7% (11.5, 52.3%). There was a significant difference between the center slice and proximal slice ($Z = -3.574$, $p < 0.001$) and between the proximal slice and distal slice ($Z = -2.722$, $p = 0.006$), but there was no significant difference between the center slice and the distal slice ($Z = -1.728$, $p = 0.084$).

Fat Fraction and Muscle Strength

Participants with BMD had statistically significant lower strength in trunk extension and flexion compared to HC (**Figure 3**). Median value of maximal trunk extension strength in participants with BMD was 88.4 Nm (48.1, 184.4 Nm) and

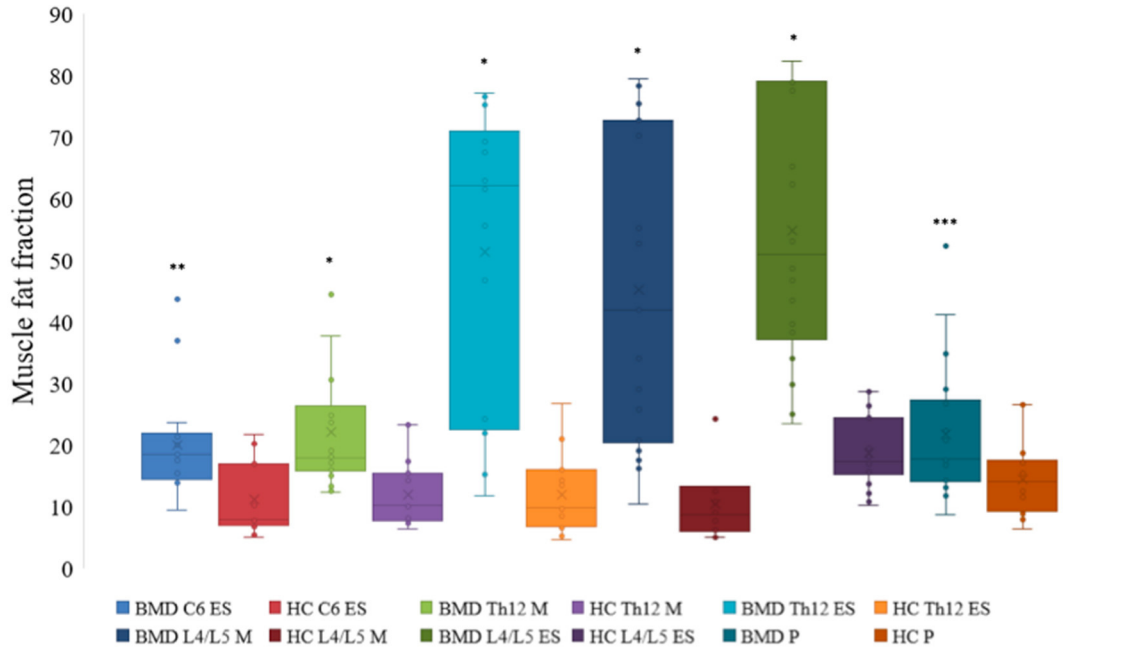


FIGURE 2 | Fat fraction of muscles at spinal level C6, Th12, and L4/L5 in Becker muscular dystrophy and healthy controls. Fat fraction of erector spinae at spinal level C6, Th12, and L4/L5, multifidus at spinal level Th12 and L4/L5, and iliopsoas of participants with Becker muscular dystrophy and healthy controls. Asterix indicate difference from healthy controls $^{*}p < 0.001$, $^{**}p = 0.001$, and $^{***}p = 0.023$. Missing value for C6 erector spinae ($n = 1$, due to phase-shift artifacts). BMD, Becker muscular dystrophy; HC, Healthy controls; ES, erector spinae; M, multifidus; P, iliopsoas.

284.6 Nm (260.9, 431 Nm) in HC. Median value of trunk flexion strength was 93.7 Nm (65.8, 134.1 Nm) in BMD and 141.1 Nm (124.3, 154.8 Nm) in HC. There was a strong and inverse correlation between global muscle FF of combined Th12 and L4/L5 spinal level muscles (multifidus and erector spinae) and trunk extension strength in participants with BMD ($\rho = -0.829$, $p < 0.001$) (**Figure 4A**). The asymptomatic participant displayed a strength of 163.1 Nm in trunk extension and 41.2 Nm in trunk flexion, and his matched healthy control displayed 434.7 Nm in extension and 146.6 in flexion.

Of the 18 BMD participants, 14 also underwent muscle strength measures of the hip and thigh. We found a strong and inverse correlation between gluteus maximus muscle FF and hip extension strength ($\rho = -0.701$, $p = 0.005$) (**Figure 4B**) but no significant relationship was found between FF of the iliopsoas muscle and hip flexion strength ($\rho = -0.279$, $p = 0.334$). A strong and inverse correlation was found between muscle FF of the knee extensor muscles (quadriceps and sartorius) and knee extension strength ($\rho = -0.842$, $p < 0.001$), and knee flexor muscles (hamstrings) and knee flexion strength ($\rho = -0.864$, $p < 0.001$) (**Figures 4C,D**).

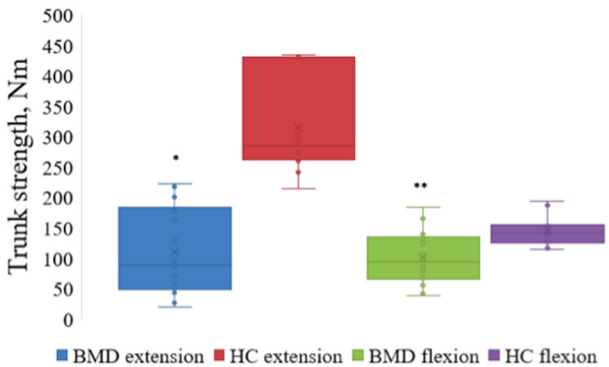


FIGURE 3 | Maximal strength of trunk extension and flexion in Becker muscular dystrophy and healthy controls. Maximal strength of trunk extension and flexion in participants with Becker muscular dystrophy and healthy controls. Asterix indicate difference from healthy controls $^{*}p < 0.001$, $^{**}p = 0.002$. Nm, Newton-Meter; BMD, Becker muscular dystrophy; HC, Healthy controls.

Correlations

The disease duration was 20.3 ± 12.7 years. There was a moderate correlation between disease duration and muscle FF of C6 erector spinae ($\rho = 0.500$, $p = 0.049$), but no correlation was found between disease duration and global muscle FF of multifidus and erector spinae at spinal level Th12 and L4/L5 (**Figure 5A**).

No correlation was found between age and muscle FF of C6 erector spinae, Th12 multifidus and erector spinae, and L4/L5 multifidus and erector spinae ($p > 0.05$) (**Figure 5B**).

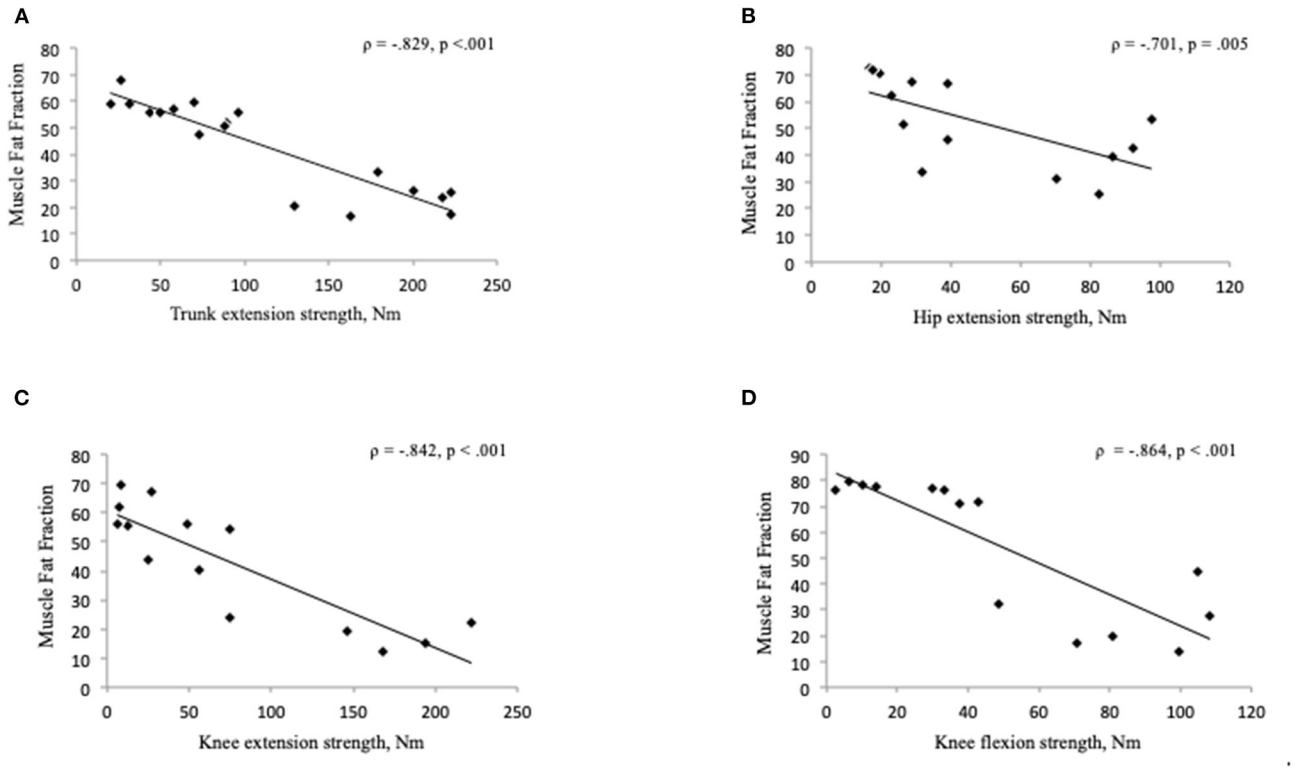


FIGURE 4 | (A–D) Correlation between muscle fat fraction and muscle strength in Becker muscular dystrophy. Correlation between global muscle fat fraction of combined Th12 and L4/L5 muscles (multifidus and erector spinae) and maximal trunk extension strength **(A)**, gluteus maximus muscle fat fraction and hip extension strength **(B)**, muscle fat fraction of the knee extensor muscles (quadriceps and sartorius) and knee extension strength **(C)**, muscle fat fraction of the knee flexor muscles (hamstrings) and knee flexion strength **(D)**. Missing value for hip and thigh strength measure ($n = 4$), due to time constraints ($n = 3$), difficulty with positioning in the Biodex chair ($n = 1$), and missing value for muscle fat fraction ($n = 4$), four data sets not used for hip and thigh because hip and thigh strength not measured on four participants. Nm, Newton-Meter.

We found a moderate correlation between muscle FF of C6 erector spinae and global muscle FF of the thigh muscles ($\rho = 0.568$, $p = 0.022$) and global muscle FF of the lower leg muscles ($\rho = 0.506$, $p = 0.046$). A strong correlation was found between global muscle FF of Th12 multifidus and erector spinae and global muscle FF of the thigh muscles ($\rho = 0.713$, $p = 0.001$), while the correlation was moderate between global muscle of Th12 multifidus and erector spinae and global muscle FF of the lower leg muscles ($\rho = 0.664$, $p = 0.004$). There was a strong correlation between global muscle FF of L4/L5 multifidus and erector spinae and global muscle FF of the thigh muscles ($\rho = 0.848$, $p < 0.001$) and global muscle FF of the lower leg muscles ($\rho = 0.725$, $p = 0.001$).

Contractile cross-sectional area of the iliopsoas muscle in relation to hip flexion strength lacked correlation ($\rho = -0.415$, $p = 0.14$) (**Figure 6A**), but not in the knee extensors (quadriceps and sartorius muscles) (**Figure 6B**) and knee flexors (hamstrings muscles) ($p < 0.05$).

There was a significant correlation between the age of the HC participants and the muscle FF levels found at spinal level C6, Th12, and L4/L5 ($p < 0.05$) (**Figure 7**).

DISCUSSION

This cross-sectional study is the first to systematically examine paraspinal muscle involvement in BMD. We studied the paraspinal muscles in a group of 18 patients with BMD using quantitative measures of trunk muscle strength and qMRI.

The main findings of our study are that paraspinal muscle FF and trunk extension strength differ significantly in patients with BMD compared to HC. Muscle FF of the paraspinal muscles is strongly and inversely correlated with trunk extension strength in BMD and muscle FF of gluteus maximus and thigh muscles correlated strongly and inversely with hip and thigh strength. Additionally, a correlation was found between muscle FF of the paraspinal muscles and muscle involvement of the thigh and lower leg muscles, indicating that disease progression of paraspinal muscles may follow general disease progression in BMD and possibly be considered as a potential biomarker.

Previous MRI studies have discovered paraspinal muscle involvement in other neuromuscular disorders. Dahlqvist et al. (14) reported paraspinal muscle affection in Facioscapulohumeral muscular dystrophy where the authors found a muscle FF of 30% at the cervical spinal level erector spinae, 45% at thoracic spinal level erector spinae, and 40%

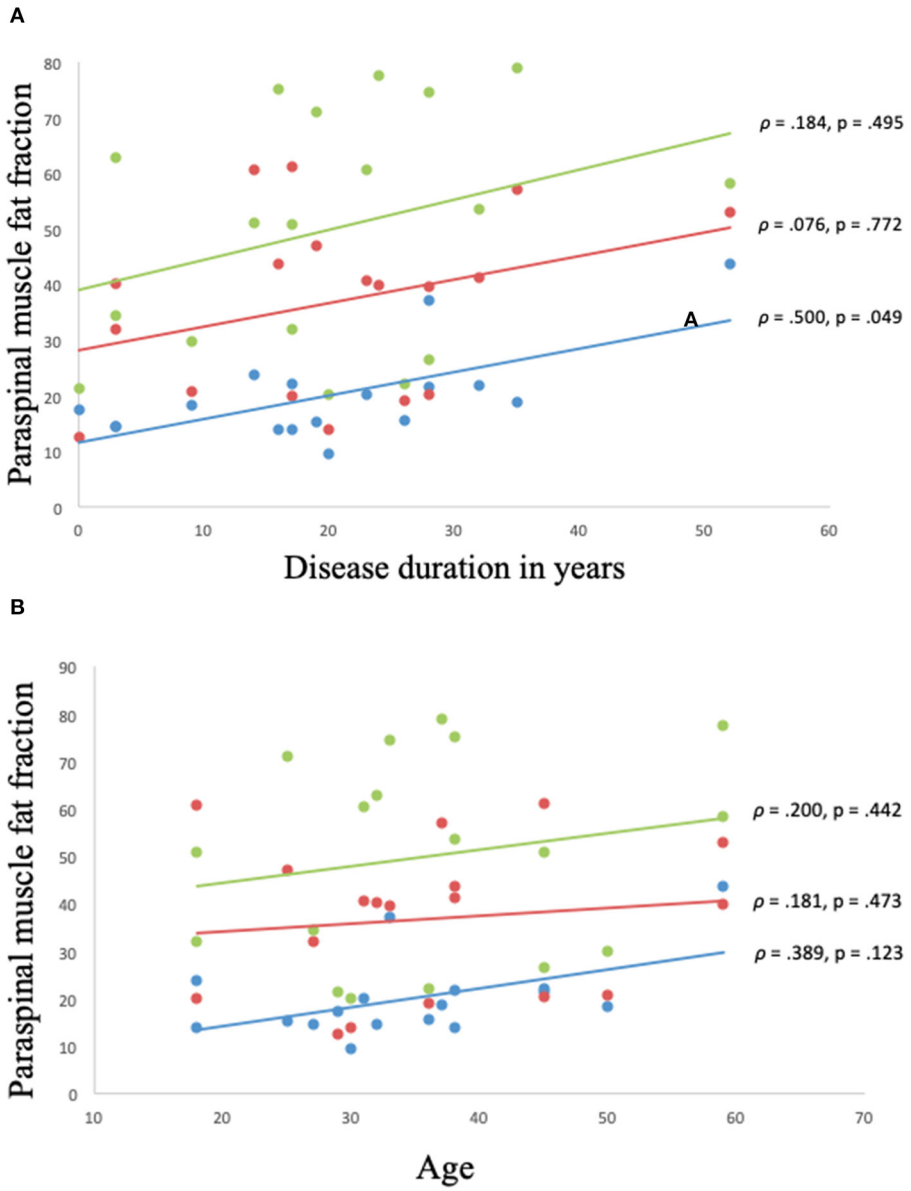


FIGURE 5 | (A,B) Correlation between muscle fat fraction of paraspinal muscles and disease duration and age in Becker muscular dystrophy. Correlation between muscle fat fraction and disease duration in Becker muscular dystrophy: Erector spinae fat fraction at C6 spinal level (blue), global muscle fat fraction at spinal level Th12 (red), and global muscle fat fraction at spinal level L4/L5 (green) **(A)**. Correlation between muscle fat fraction and age in Becker muscular dystrophy: Erector spinae fat fraction at C6 spinal level (blue), global muscle fat fraction at spinal Th12 (red), and global muscle fat fraction at spinal level L4/L5 (green) **(B)**. Missing value for **(A,B)** ($n = 1$, due to phase-shift artifacts).

at the lumbar spinal level erector spinae. In our cohort, we found a lesser muscle affection in erector spinae at the cervical region and a larger affection in multifidus and erector spinae at the thoracic and lumbar region in comparison with Facioscapulohumeral muscular dystrophy. Furthermore, a case study reported paraspinal muscle affection in a patient with McArdle disease (21), and paraspinal muscle involvement was reported in Myotonic muscular dystrophy with large fat replacement observed in erector spinae at the lumbar spinal level (22). Muscle FF of 49.9% was observed in erector spinae at the lumbar spinal level in Duchenne muscular dystrophy (23), and Schreckenbach et al. (24) found paraspinal muscle involvement in two individuals with reducing body myopathy. Frequent involvement of the paraspinal muscles was also found in female carriers of dystrophinopathy (25), underlining the importance of MRI to identify muscle pathology in muscle diseases.

Paraspinal muscles are vital for the stabilization and mobilization of the trunk and consequently, weakness, and

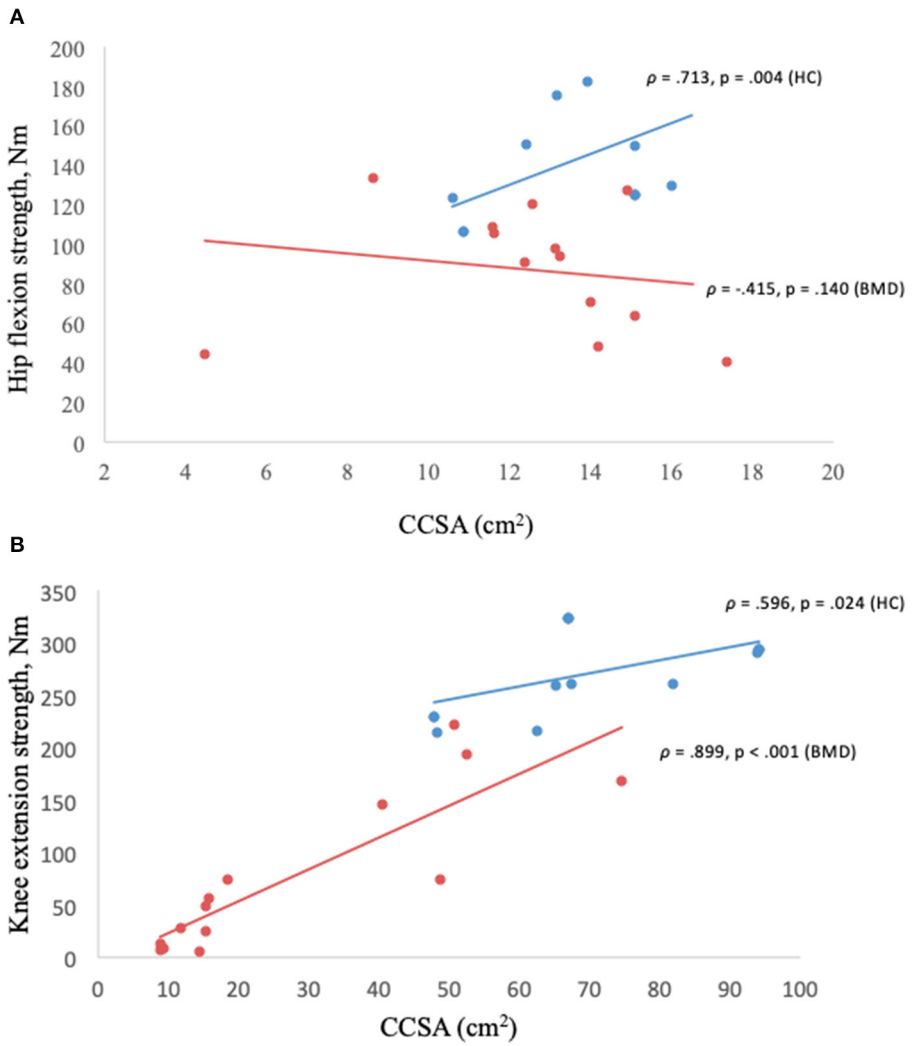


FIGURE 6 | (A,B) Muscle strength in relation to contractile cross-sectional area. Hip flexion strength and contractile cross-sectional area of iliopsoas muscle in Becker muscular dystrophy and healthy controls **(A)**, knee extension strength and contractile cross-sectional area of knee extensors (quadriceps and sartorius) in Becker muscular dystrophy and healthy controls **(B)**. Missing value for Becker muscular dystrophy ($n = 3$) due to participant time constraints and ($n = 1$) due to positioning difficulty in the scanner, missing value for healthy controls ($n = 4$) to match sample size of Becker muscular dystrophy participants. BMD, Becker muscular dystrophy; HC, Healthy controls; CCSA, contractile cross-sectional area; Nm, Newton-meter.

impairment of these muscles can have a profound impact on activities of daily living, resulting in a backward bent posture, poor mobility, postural control difficulties, and potential spinal malalignment. We assessed trunk strength using Biodex Dual-position back Extension/Flexion attachment unit. Biodex is a consistent method to assess muscle strength and has shown to be a reliable method to measure trunk muscle strength (26). Muscle FF found in the paraspinal muscles was strongly and inversely associated with reduced muscle strength in trunk extension in participants with BMD. A similar association between trunk extension strength and muscle FF was demonstrated by Dahlqvist et al. (14) where the authors found a connection between paraspinal muscle FF and back extension strength in Facioscapulohumeral muscular dystrophy. Furthermore, Schlaeger et al. (27) observed a relationship between muscle FF of erector spinae and back extension strength in healthy participants and found a muscle FF of 9%. In our study, the healthy population had a muscle FF of about 20%. The higher level of FF in our HC participants could be explained by age, in that age ranged between 18 and 59 years and it has previously been shown that paraspinal muscles are more susceptible to fat infiltration with aging (15), and Hadar et al. (28) also found a higher level of fat replacement in older males. Additionally, age in our HC group was related to muscle FF of the paraspinal muscles, despite a small sample size (**Figure 7**). In addition, Schlaeger et al. combined erector spinae and multifidus when determining the cross-sectional area, whereas, we separated the two muscles in our study. However, a combined muscle FF of erector spinae

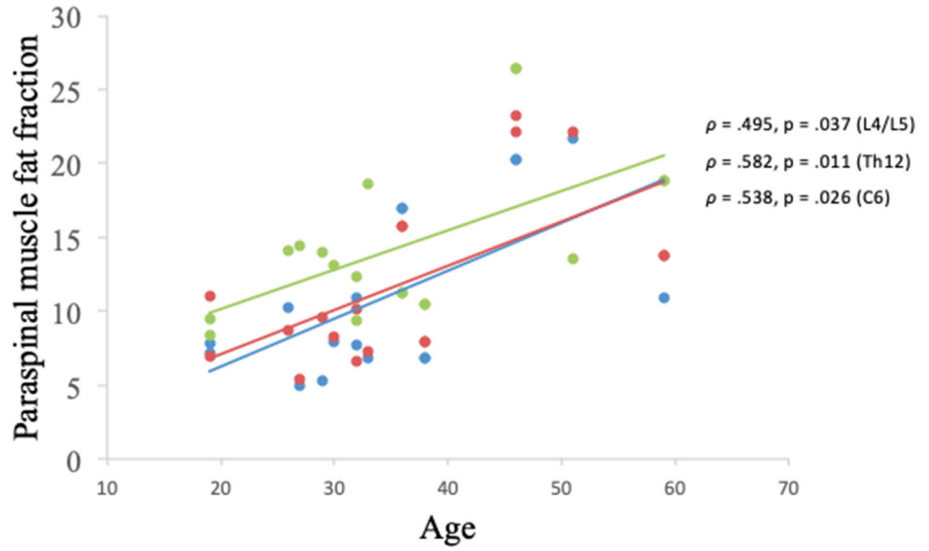


FIGURE 7 | Correlation between muscle fat fraction of paraspinal muscles and age in healthy controls. Erector spinae muscle fat fraction at C6 spinal level (blue), global muscle fat fraction at spinal level Th12 (red), and global muscle fat fraction at spinal level L4/L5 (green).

and multifidus resulted in a muscle FF of about 15% in our HC participants. Furthermore, Dahlqvist et al. (14) found an FF of 20% in their healthy population, which corresponds well with our study. Moreover, a previous study found a connection between fat replacement in the calf muscle and weakness in plantar flexors in patients with BMD, amplifying the relationship between muscle FF and muscle strength (20).

There was no correlation between disease duration and muscle FF of the paraspinal muscles, except at spinal level C6 erector spinae, and age, which could be explained by the heterogeneity of the clinical phenotype of BMD. In addition, we found a difference in muscle FF distribution in the quadriceps muscle and calf muscle, suggesting that muscle FF distribution is not homogenous in these muscles. Furthermore, muscle involvement is less severe in the proximal part of the quadriceps and calf muscles in relation to the center and distal slice of the muscles.

We found a large muscle FF in the gluteus maximus muscle, corresponding well to previous studies where a severe involvement of the gluteus maximus in BMD was established (13, 17, 29). The iliopsoas muscle was less affected in participants with BMD relative to the muscles of the thoracic spine, lumbar spine, and lower limbs. This finding agrees with an earlier study where the authors found that iliopsoas was less affected in patients with BMD relative to lower limb muscles (13). In other neuromuscular disorders, however, iliopsoas is prominently affected, such as in Bethlem myopathy (30), Hypokalemic Periodic Paralysis (31), and Limb-Girdle muscular dystrophy type R9 (32). The difference in the level of muscle affection in iliopsoas and paraspinal muscles may be linked to the specific molecular defect, but also variability in muscle fiber composition of these muscles. Earlier studies have shown that iliopsoas has a predominance of type 2A muscle fibers (33) and paraspinal muscles have a predominance of type 1 fibers (34). However, whether muscle fiber composition has an impact on the level of muscle involvement in BMD is unknown. Furthermore, our study showed that contractile properties of the iliopsoas muscle in relation to hip flexion strength were disrupted. A previous study also reported a contractile disruption of the calf muscle in BMD (20). In Duchenne muscular dystrophy, Wokke et al. (35) found a correlation between CSAA and muscle strength in the quadriceps muscle and hamstring muscles which corresponds well with our findings in the quadriceps and hamstring muscles in BMD. Additionally, Wokke et al. (35) found a disruption in contractile properties in the triceps surae muscle, resembling the findings of Løkken et al. (20). This suggests that there may be similarities to be found in the contractile properties in BMD and Duchenne muscular dystrophy. Disrupted contractile properties have also been found in other muscle disorders such as congenital myopathy (36) and spinobulbar muscular atrophy (37). Knee extension muscle strength ranged between 6.1 and 222 Nm in BMD and in HC the strength ranged between 214.8 and 323.9 Nm, indicating that patients with BMD have about 31–97% reduced muscle strength than HC. Despite a substantially lower muscle strength in knee extension, contractile properties in the quadriceps and sartorius muscles were not disrupted in BMD, suggesting that contractile properties may be related to muscle type, muscle fiber type composition, and the level of severity of muscle involvement.

Imaging studies using visual rating scales in BMD have distinctively been able to show patterns of muscle involvement (13, 29). However, to quantify pathological changes in the muscle and to monitor disease progression, qMRI has been shown to be more sensitive than a visual rating scale (8, 19). Quantitative MRI represents a valid biomarker

and provides an objective endpoint to measure disease distribution (38). Intramuscular changes observed in the asymptomatic BMD participant show that MRI can detect subtle intramuscular degenerative changes before symptoms occur. Furthermore, previous studies done in patients with BMD have presented a correlation between muscle FF in the lower limbs and functional measures such as 6-min walk test, motor function measure, and the north star ambulatory assessment (7, 8, 19), suggesting that qMRI is valuable in the evaluation of the disease as a complementary tool to clinical functional testing.

Our study has some limitations: The investigated cohort of patients is relatively small. However, we were able to determine a relatively consistent muscle involvement in the paraspinal muscles. We did not test isokinetic muscle strength because the weight of the trunk flexion and extension unit exceeded what the majority of the patients could manage, and therefore, it is unknown whether isokinetic trunk strength assessment provides similar results. However, isometric testing is a reliable method to identify maximal strength in a muscle and has been shown to have a high intra-class correlation in relation to isokinetic testing (39). A 2-point Dixon technique may be more sensitive to phase-shift artifacts than a 3-point Dixon technique (40); However, we experienced phase-shift in just one participant at spinal level C6 and that one image was not included in the analysis.

We conclude that patients with BMD demonstrate severe paraspinal muscle involvement. The level of muscle involvement is based on fat replacement within the muscle and decreased muscle strength. We found high levels of fat replacement of the paraspinal muscles which strongly and inversely correlates with reduced muscle strength in trunk extension and the level of FF in the paraspinal muscles parallel involvement of the lower limb muscles. Findings from this study contribute to the clinical management of patients with BMD and suggest that assessment of paraspinal muscle strength, and fat replacement may serve as a possible biomarker in longitudinal investigations of BMD.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Danish National Committee on Health Research Ethics (approval number: H-16030358). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the participants for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

AMS contributed with data collection, data analysis, and drafting of manuscript. KR contributed with study design, data collection, and drafting of manuscript. JS contributed with data collection. TK contributed with data collection. NW contributed with study design and drafting of manuscript. JV contributed with study design, data analysis, and drafting of manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Thanks to Grosserer L. F. Foghts Fond, Oda og Hans fond, Toyota-Fonden, Aase og Ejnar Danielsens Fond, Torkild Steenbecks Legat, Torben and Alice Frimodts Fond, and Hede Nielsens Fond for providing grants.

ACKNOWLEDGMENTS

We thank the patients for their participation in this study.

REFERENCES

1. Kendall FP. *Muscles Testing and Function with Posture and Pain*. 5th ed. Philedelphia, PA: Lippincott Willioams & Wilkins (2005).
2. Witting N, Andersen LK, Vissing J. Axial myopathy: an overlooked feature of muscle diseases. *Brain J Neurol*. (2016) 139(Pt. 1):13–22. doi: 10.1093/brain/awv332
3. Moreland J, Finch E, Stratford P, Balsor B, Gill C. Interrater reliability of six tests of trunk muscle function and endurance. *J Orthop Sports Phys Ther*. (1997) 26:200–8. doi: 10.2519/jospt.1997.26.4.200
4. Gaeta M, Scribano E, Mileto A, Mazziotti S, Rodolico C, Toscano A, et al. Muscle fat fraction in neuromuscular disorders: dual-echo dual-flip-angle spoiled gradient-recalled MR imaging technique for quantification–a feasibility study. *Radiology*. (2011) 259:487–94. doi: 10.1148/radiol.10101108
5. Lamminen AE. Magnetic resonance imaging of primary skeletal muscle diseases: patterns of distribution and severity of involvement. *Br J Radiol*. (1990) 63:946–50. doi: 10.1259/0007-1285-63-756-946
6. Wattjes MP, Kley RA, Fischer D. Neuromuscular imaging in inherited muscle diseases. *Eur Radiol*. (2010) 20:2447–60. doi: 10.1007/s00330-010-1799-2
7. Fischer D, Hafner P, Rubino D, Schmid M, Neuhaus C, Jung H, et al. The 6-minute walk test, motor function measure and quantitative thigh muscle MRI in Becker muscular dystrophy: a cross-sectional study. *Neuromuscul Disord*. (2016) 26:414–22. doi: 10.1016/j.nmd.2016.04.009
8. Maggi L, Moscatelli M, Frangiamore R, Mazzi F, Verri M, De Luca A, et al. Quantitative muscle MRI protocol as possible biomarker in Becker muscular dystrophy. *Clin Neuroradiol*. (2020) 31:257–66. doi: 10.1007/s00062-019-00875-0
9. Morrow JM, Sinclair CDJ, Fischmann A, Machado PM, Reilly MM, Yousry TA, et al. MRI biomarker assessment of neuromuscular disease progression: a prospective observational cohort study. *Lancet Neurol*. (2016) 15:65–77. doi: 10.1016/S1474-4422(15)00242-2
10. Willis TA, Hollingsworth KG, Coombs A, Sveen M-L, Andersen S, Stojkovic T, et al. Quantitative muscle MRI as an assessment tool for monitoring disease progression in LGMD2I: a multicentre longitudinal study. *PLoS ONE*. (2013) 8:e70993. doi: 10.1371/journal.pone.0070993
11. Andersen G, Dahlqvist JR, Vissing CR, Heje K, Thomsen C, Vissing J. MRI as outcome measure in facioscapulohumeral muscular dystrophy: 1-year follow-up of 45 patients. *J Neurol. March.* (2017) 264:438–47. doi: 10.1007/s00415-016-8361-3
12. Dahlqvist JR, Fornander F, Borch J de S, Oestergaard ST, Poulsen NS, Vissing J. Disease progression and outcome measures in spinobulbar muscular atrophy. *Ann Neurol.* (2018) 84:754–65. doi: 10.1002/ana.25345

13. Tasca G, Iannaccone E, Monforte M, Masciullo M, Bianco F, Laschena F, et al. Muscle MRI in Becker muscular dystrophy. *Neuromuscul Disord.* (2012). 22 (Suppl. 2):S100–6. doi: 10.1016/j.nmd.2012.05.015
14. Dahlqvist JR, Vissing CR, Thomsen C, Vissing J. Severe paraspinal muscle involvement in facioscapulohumeral muscular dystrophy. *Neurology*. (2014) 83:1178–83. doi: 10.1212/WNL.0000000000000828
15. Dahlqvist JR, Vissing CR, Hedermann G, Thomsen C, Vissing J. Fat replacement of paraspinal muscles with aging in healthy adults. *Med Sci Sports Exerc.* (2017) 49:595–601. doi: 10.1249/MSS.0000000000001119
16. Dahlqvist JR, Oestergaard ST, Poulsen NS, Thomsen C, Vissing J. Refining the spinobulbar muscular atrophy phenotype by quantitative MRI and clinical assessments. *Neurology*. (2019) 92:e548–59. doi: 10.1212/WNL.0000000000006887
17. Barp A, Bello L, Caumo L, Campadello P, Semplicini C, Lazzarotto A, et al. Muscle MRI and functional outcome measures in Becker muscular dystrophy. *Sci Rep*. (2017) 7:16060. doi: 10.1038/s41598-017-16170-2
18. Sheikh AM, Rudolf K, Witting N, Vissing J. Quantitative muscle MRI as outcome measure in patients with becker muscular dystrophy-A 1-year follow-up study. *Front Neurol.* (2021) 11:613489. doi: 10.3389/fneur.2020.613489
19. Bonati U, Schmid M, Hafner P, Haas T, Bieri O, Gloor M, et al. Longitudinal 2-point dixon muscle magnetic resonance imaging in becker muscular dystrophy. *Muscle Nerve*. (2015) 51:918–21. doi: 10.1002/mus.24629
20. Løkken N, Hedermann G, Thomsen C, Vissing J. Contractile properties are disrupted in Becker muscular dystrophy,s but not in limb girdle type 2I. *Ann Neurol*. (2016) 80:466–71. doi: 10.1002/ana.24743
21. Witting N, Duno M, Piraud M, Vissing J. Severe axial myopathy in McArdle disease. *JAMA Neurol.* (2014) 71:88–90. doi: 10.1001/jamaneurol.2013.3209
22. Solbakken G, Bjørnarå B, Kirkhus E, Nguyen B, Hansen G, Frich JC, et al. MRI of trunk muscles and motor and respiratory function in patients with myotonic dystrophy type 1. *BMC Neurol*. (2019) 19:135. doi: 10.1186/s12883-019-1357-8
23. Zoabli G, Mathieu PA, Aubin C-E. Magnetic resonance imaging of the erector spinae muscles in Duchenne muscular dystrophy: implication for scoliotic deformities. *Scoliosis*. (2008) 3:21. doi: 10.1186/1748-7161-3-21
24. Schreckenbach T, Henn W, Kress W, Roos A, Maschke M, Feiden W, et al. Novel FHL1 mutation in a family with reducing body myopathy. *Muscle Nerve*. (2013) 47:127–34. doi: 10.1002/mus.23500
25. Tasca G, Monforte M, Lannaccone E, Laschena F, Ottaviani P, Silvestri G, et al. Muscle MRI in female carriers of dystrophinopathy. *Eur J Neurol.* (2012) 19:1256–60. doi: 10.1111/j.1468-1331.2012.03753.x
26. García-Vaquero MP, Barbado D, Juan-Recio C, López-Valenciano A, Vera-Garcia JF. Isokinetic trunk flexion-extension protocol to assess trunk muscle strength and endurance: reliability, learning effect, and sex differences. *J Sport Health Sci.* (2020) 9:692–701. doi: 10.1016/j.jshs.2016.08.011
27. Schlaeger S, Inhuber S, Rohrmeier A, Dieckmeyer M, Freitag F, Klupp E, et al. Association of paraspinal muscle water-fat MRI-based measurements with isometric strength measurements. *Eur Radiol*. (2019) 29:599–608. doi: 10.1007/s00330-018-5631-8
28. Hadar H, Gadoth N, Heifetz M. Fatty replacement of lower paraspinal muscles: normal and neuromuscular disorders. *Am J Roentgenol*. (1983) 141:895–8. doi: 10.2214/ajr.141.5.895
29. Faridian-Aragh N, Wagner KR, Leung DG, Carrino JA. Magnetic resonance imaging phenotyping of Becker muscular dystrophy. *Muscle Nerve*. (2014) 50:962–7. doi: 10.1002/mus.24246
30. Salim R, Dahlqvist JR, Khawajazada T, Kass K, Revsbech KL, de Stricker Borch J, et al. Characteristic muscle signatures assessed by quantitative MRI in patients with Bethlem myopathy. *J Neurol*. (2020) 267:2432–42. doi: 10.1007/s00415-020-09860-x
31. Holm-Yildiz S, Witting N, Dahlqvist J, de Stricker Borch J, Solheim T, Fornander F, et al. Permanent muscle weakness in hypokalemic periodic paralysis. *Neurology*. (2020) 95:e342–52. doi: 10.1212/WNL.0000000000009828
32. Revsbech K, Khawajazada T, Borch J, Rudolf K, Sheikh A, Dahlqvist J, et al. Vissing. P.180 Paraspinal muscle affection in limb-girdle muscular dystrophy type 2I patients. *Neuromusc Disord.* (2019) 29(Supplement 1):S101. doi: 10.1016/j.nmd.2019.06.235
33. Arbanas J, Klasan GS, Nikolic M, Jerkovic R, Miljanovic I, Malnar D. Fibre type composition of the human psoas major muscle with regard to the level of its origin. *J Anat*. (2009) 215:636–41. doi: 10.1111/j.1469-7580.2009.01155.x
34. Sirca A, Kostevc V. The fibre type composition of thoracic and lumbar paravertebral muscles in man. *J Anat*. (1985) 141:131–7.
35. Wokke BH, van den Bergen JC, Versluis MJ, Niks EH, Milles J, Webb AG, et al. Quantitative MRI strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. *Neuromuscul Disord*. (2014) 24:409–16. doi: 10.1016/j.nmd.2014.01.015
36. Eisum A-SV, Fornander F, Poulsen NS, Andersen AG, Dahlqvist J, Andersen LK, et al. Contractile properties are impaired in congenital myopathies. *Neuromuscul Disord*. (2020) 30:649–55. doi: 10.1016/j.nmd.2020.06.007
37. Dahlqvist JR, Oestergaard ST, Poulsen NS, Knak KL, Thomsen C, Vissing J. Muscle contractility in spinobulbar muscular atrophy. *Sci Rep*. (2019) 9:4680. doi: 10.1038/s41598-019-41240-y
38. Dahlqvist JR, Widholm P, Leinhard OD, Vissing J. MRI in neuromuscular diseases: an emerging diagnostic tool and biomarker for prognosis and efficacy. *Ann Neurol.* (2020) 88:669–81. doi: 10.1002/ana.25804
39. Adsuar JC, Olivares PR, del Pozo-Cruz B, Parraca JA, Gusi N. Test-retest reliability of isometric and isokinetic knee extension and flexion in patients with fibromyalgia: evaluation of the smallest real difference. *Arch Phys Med Rehabil.* (2011) 92:1646–51. doi: 10.1016/j.apmr.2011.04.017
40. Coombs BD, Szumowski J, Coshow W. Two-point Dixon technique for water-fat signal decomposition withB0 inhomogeneity correction. *Magn Reson Med*. (1997) 38:884–9. doi: 10.1002/mrm.1910380606

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 17 June 2021
doi: 10.3389/fneur.2021.636719

Use of EP3533-Enhanced Magnetic Resonance Imaging as a Measure of Disease Progression in Skeletal Muscle of *mdx* Mice

Alexander Peter Murphy[1,2], *Elizabeth Greally*[2], *Dara O'Hogain*[3], *Andrew Blamire*[3], *Peter Caravan*[4] *and Volker Straub*[1,5]*

[1] *The Institute of Cancer and Genomics, Birmingham University, Birmingham, United Kingdom,* [2] *The John Walton Muscular Dystrophy Research Centre, Institute of Translational and Clinical Research, Newcastle University, Newcastle upon Tyne, United Kingdom,* [3] *Newcastle Magnetic Resonance Centre, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom,* [4] *Department of Radiology, Martinos Center for Biomedical Imaging, Harvard Medical School, Massachusetts General Hospital, Charlestown, MA, United States,* [5] *Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle University, Newcastle upon Tyne, United Kingdom*

OPEN ACCESS

Edited by:
Anna Pichiecchio, Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy

Reviewed by:
Jill Rafael-Fortney, The Ohio State University, United States
Russell G. Rogers, Cedars-Sinai Medical Center, United States
Heinrich Brinkmeier, Universitätsmedizin Greifswald, Germany

***Correspondence:**
Volker Straub
Volker.straub@newcastle.ac.uk

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: *01 December 2020*
Accepted: *12 May 2021*
Published: *17 June 2021*

Citation:
Murphy AP, Greally E, O'Hogain D, Blamire A, Caravan P and Straub V (2021) Use of EP3533-Enhanced Magnetic Resonance Imaging as a Measure of Disease Progression in Skeletal Muscle of mdx Mice. Front. Neurol. 12:636719. doi: 10.3389/fneur.2021.636719

As putative treatments are developed for Duchenne muscular dystrophy (DMD), sensitive, non-invasive measures are increasingly important to quantify disease progression. Fibrosis is one of the histological hallmarks of muscular dystrophy and has been directly linked to prognosis. EP3533 is a novel contrast agent with an affinity to collagen 1 that has demonstrated a significant and high correlation to *ex vivo* fibrosis quantification. Halofuginone is an established anti-fibrotic compound shown to reduce collagen skeletal muscle fibrosis in murine models of DMD. This experiment explored whether EP3533 could be used to detect signal change in skeletal muscle of *mdx* mice before and after a 12 week course of halofuginone compared to controls. Four age-matched groups of treated and untreated mice were evaluated: 2 groups of *mdx* ($n = 8$ and $n = 13$, respectively), and 2 groups of BL10 mice ($n = 5$ and $n = 3$, respectively). Treated mice received an intraperitoneal injection with halofuginone three times per week for 12 weeks, with the remaining mice being given vehicle. Both *mdx* groups and the untreated BL10 were scanned at baseline, then all groups were scanned on week 13. All subjects were scanned using a 7T Varian scanner before and after administration of EP3533 using a T1 mapping technique. Mice underwent grip testing in week 13 prior to dissection. Skeletal muscle was used for Masson's trichrome quantification, hydroxyproline assay, and immunofluorescent antibody staining. Untreated *mdx* mice demonstrated a significant increase in R1 signal from pre- to post-treatment scan in three out of four muscles (gastrocnemius $p = 0.04$, hamstrings $p = 0.009$, and tibialis anterior $p = 0.01$), which was not seen in either the treated *mdx* or the BL10 groups. Histological quantification of fibrosis also demonstrated significantly higher levels in the untreated *mdx* mice with significant correlation seen between histology and EP3533 signal change. Forelimb weight adjusted-grip strength was significantly lower in the untreated *mdx* group, compared to the treated group. EP3533 can be used over time as an outcome measure to quantify treatment effect of an established anti-fibrotic drug. Further studies are needed to evaluate the use of this contrast agent in humans.

Keywords: EP3533, muscular dystrophy, fibrosis, magnetic resonance image, *mdx* mouse model

INTRODUCTION

Duchenne muscular dystrophy (DMD) is caused by mutations in the *DMD* gene and is inherited in an X-linked recessive fashion. DMD is considered the most commonly inherited muscle disease in childhood, with an incidence of 1 in 3,500–6,000 live male births (1–3). Dystrophin, the protein product of the *DMD* gene, is a sub-sarcolemmal cytoskeletal protein present in all muscle (4). The phenotype of patients with DMD is of progressive weakness of the skeletal, respiratory and cardiac muscles. DMD is considered a devastating, life-limiting condition (3, 5).

Dystrophin-deficient muscle results in a loss of structural membrane integrity of muscle fibres during cycles of contraction and relaxation. The damaged sarcolemma leads to an influx of cations such as calcium and sodium (6) into the intracellular compartment and consecutively to myofiber necrosis and chronic inflammation. Subsequently there is an accumulation of fibrosis and adipose tissue within the muscle (7). Higher degrees of endomysial fibrosis in skeletal muscle has been associated with a worse prognosis in neuromuscular disease (8). Sensitive and preferably minimally invasive techniques are increasingly important to evaluate putative therapies by quantifying fibrosis *in vivo* as a biomarker of disease progression.

EP3533 is a gadolinium-based MRI contrast agent which has an affinity for type 1 collagen (9). EP3533 has been successful in quantifying and staging fibrosis in several murine models (9–13) as well as quantifying anti-fibrotic therapy response (14). EP3533-enhanced MRI has been previously evaluated in a murine model of DMD, quantifying fibrosis in skeletal and cardiac tissues (15). There was a significant linear correlation demonstrated between *ex vivo* measures of fibrosis and skeletal muscle measurement of R1 change using EP3533-enhanced MRI (gastrocnemius: $r = 0.83$, $P = 0.001$; tibialis anterior: $r = 0.73$, $P = 0.01$) (15). While there is evidence that EP3533-enhanced MRI could be used to quantify fibrosis, it is unclear whether it is sensitive enough to detect disease progression in *mdx* mice and able to demonstrate changes in response to treatment as an outcome measure.

Halofuginone is a potent anti-fibrotic which inhibits phosphorylation of Smad 3 and attenuates gene expression of collagen-α1 in fibroblasts. These effects have been shown to lead to a reduction in collagen deposition and fibrosis (16, 17). Halofuginone has been trialled as an anti-fibrotic in the treatment of *mdx* mice with success *in vivo* (17–19). Studies have shown that halofuginone administration to young and old *mdx* mice have demonstrated significant improvements in measures of fibrosis and skeletal muscle function (18, 19).

The primary aim of this study was to determine whether EP3533-enhanced quantitative MRI could be used to demonstrate significant differences between *mdx* mice treated with an established anti-fibrotic (halofuginone). The secondary aim was to evaluate whether EP3533-enhanced quantitative MRI correlated significantly to *ex vivo* measures of fibrosis and a measurement of muscle function.

MATERIALS AND METHODS

Four groups of age-matched (16 ± 3 weeks) *mdx* (C57BL/10ScSn-mdx/J, Jackson, Maine USA)and BL10 (C57BL/10ScSnOlaHsd, wild type, WT, Harlan Laboratories, Indianapolis, USA) male mice were utilised for the study. Groups were split to receive either vehicle (150 μl of 5% dimethyl sulfoxide and 0.9% saline) or halofuginone (7.5 μg) dissolved in the same solution (18, 20). All groups received vehicle or halofuginone via intraperitoneal injections three times per week for 12 weeks. The treated *mdx* group ($n = 8$), untreated *mdx* group ($n = 13$) and the treated BL10 mice ($n = 3$) would undergo baseline MRI scans with EP3533. The untreated BL10 group ($n = 5$) would be scanned at follow up (week 13) only. All groups underwent grip strength testing at week 13, then scanning of the lower limbs. Following this, mice were humanely killed and dissected for *ex vivo* quantification of fibrosis.

MRI Protocol

Animals were anaesthetised using 5% isoflurane in 0.5 L/min of oxygen, once induced, anaesthesia was maintained at 1 to 2% isoflurane. Hair removal cream was used to clear two patches of skin for contact with ECG electrodes, the tail vein cannulated and a line containing the EP3533 attached. The mouse was placed on a sled, incorporating ECG electrodes for cardiac imaging (Dazai, Canada) and positioned in a 39 mm internal diameter birdcage radiofrequency coil (Rapid Biomedical GmbH, Germany). Core body temperature was maintained using a warm air heating system interfaced to a rectal temperature monitoring probe and heart rate and respiration monitored (SA Instruments, NY).

Magnetic resonance scanning was performed on a 7-Tesla micro-imaging system (Varian; Agilent Technologies, Santa Clara, United States) equipped with a 12 cm micro-imaging gradient insert (maximum gradient, 40 mT/m). Following administration of 20 μmol/kg of EP3533, based on a previous study (15), subjects underwent all imaging at 60–70 min post-infusion using T1 weighted gradient recalled echo (GRE) imaging and then a gradient echo multi-slice inversion-recovery Look-Locker (gemsIR-LL). The T1 mapping sequence was performed on both upper and lower portions of the rear legs using a gemsIR-LL with 1 slice of 1 mm thickness selected through lower leg muscles with the following parameters: repetition time (TR) inversion = 5 s, 10 inversion times (from 0.1 to 5 s), TR/echo time (TE) = 9.68/4.86 ms, echo train length (ETL) = 4, flip angle (FA) = 4°, Field Of View (FOV) 30 × 30 mm, matrix = 128 × 128. The T1 weighted GRE acquired pre and post-contrast enhancement was acquired with TR/TE = 9.46/4.75 ms, FA = 20°, FOV = 30 × 30 × 30 mm, matrix 128 × 128 × 128, band width = 20 KHz.

For the lower leg muscles, mice were positioned with the upper border of the field of view beginning at the lower border of the knee. For the upper leg, mice were positioned with the centre of the knee joint as close as possible to the centre of the radiofrequency coil. Four slices were used to obtain the gemsIR-LL, two positioned above (+0.6 mm, 1.2 mm) and two below the knee joint (−0.6 mm, −1.2 mm). This enabled homogeneity in positioning for analysis.

All procedures performed were in accordance with the ethical standards of directive 2010/63/EU of the European parliament and under the auspices of the terms of the animals (scientific procedures) act 1986 and project licence PB3CA650C, authorised by the Home Secretary, Home Office, United Kingdom.

MRI Analysis and Region of Interest Selection

All scans were analysed using "Aedes software" (21). T1 maps were created by pixel-wise fitting of the lower limb datasets to the standard inversion recovery Look-Locker equation. Lower limb muscle ROIs were interactively defined in tibialis anterior and gastrocnemius muscles, based on previous MRI-based murine studies (22). Application of the ROIs to T1 maps were used to determine the change in R1 values in skeletal and cardiac muscles pre and post-contrast.

Ex vivo Fibrosis Quantification

Mice were humanely killed immediately after the second scan. The tibialis anterior, gastrocnemius, hamstring and quadriceps muscles were harvested. Samples were snap frozen in liquid nitrogen cooled isopentane and stored at −80°C. The left-sided muscles were collected for quantification of fibrosis markers via hydroxyproline, while right-sided muscles were obtained for histological analysis. In all muscles cryosections were cut 8 μm thick. To ensure more coverage of the muscle to detect focal fibrosis, sections were spaced 200 μm apart.

Histological staining was used to quantify fibrosis and confirm disease pathology. Cryosections were stained with Masson's trichrome (MaTr), haematoxylin and eosin (H&E) and immunofluorescent staining for collagen 1.

The percentage of fibrosis was quantified using the MaTr staining and regions of interest (ROIs) analysis via imaging software (Fiji 64 bit). Fibrosis was calculated as a percentage of the total area of the muscle section. A modified hydroxyproline assay was also used to quantify fibrosis in muscle (23).

Pharmacokinetics Sampling

To confirm serum levels of halofuginone, pharmacokinetics (PK) was performed in four halofuginone-naive *mdx* mice of similar ages to the experimental groups, using liquid chromatography—mass spectrometry/mass spectrometry (LC-MS/MS). Venesection of the saphenous vein was performed using a needle and capillary tubes after administration of halofuginone. To keep within UK Home Office guidance for volume of venesection per mouse, the samples were taken from four mice at two different time points each (24). From each mouse two samples of 150 μl were taken, with the second a terminal sample. PK levels were measured at: pre-injection, 5, 15, and at 30 min, then at 1, 2, 4, 8, and 24 h. These were stored as serum in lithium-heparin bottles and stored at −80°C. LC-MS/MS allowed quantification against a calibration curve from 0.5 to 1000 ng/ml. A stock solution of halofuginone in DMSO (5%) and saline (0.9%) was provided at a concentration of 10 mM

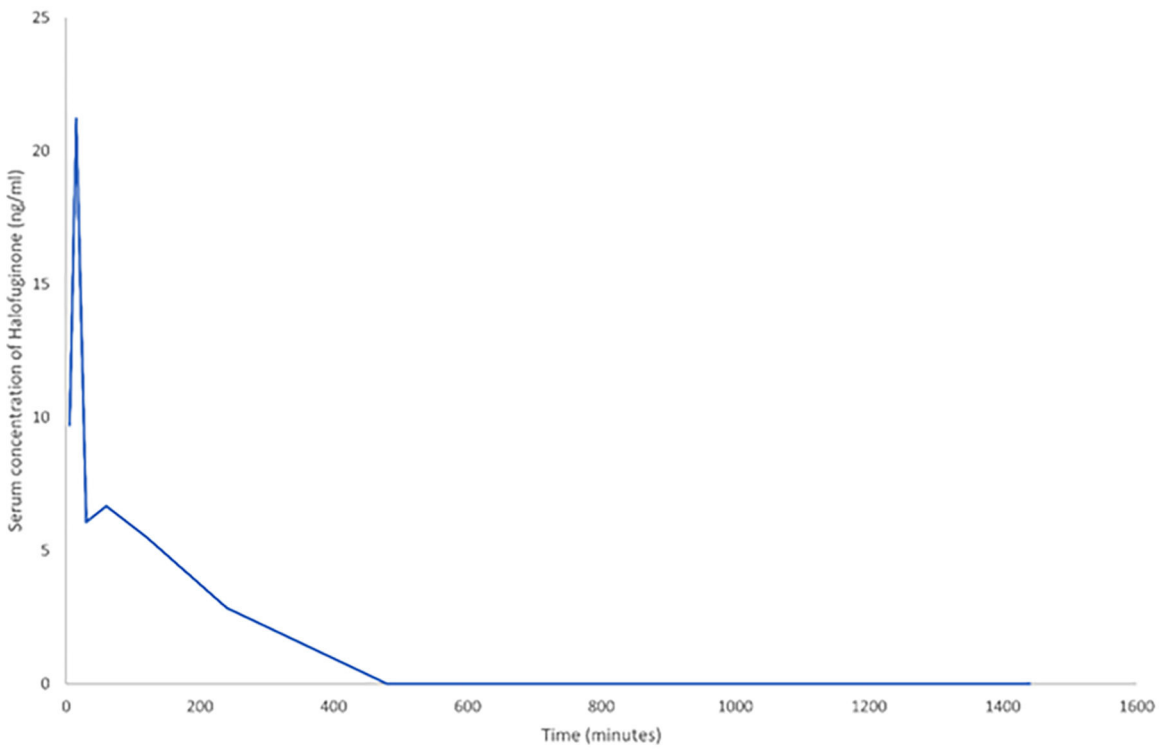


FIGURE 1 | A line graph demonstrating the change in serum halofuginone post-intraperitoneal (IP) administration of 7.5 μg ($n = 4$). To allow inclusion of all time points, time is measured in minutes. The last five data points are at 1, 2, 4, 8, and 24 h respectively.

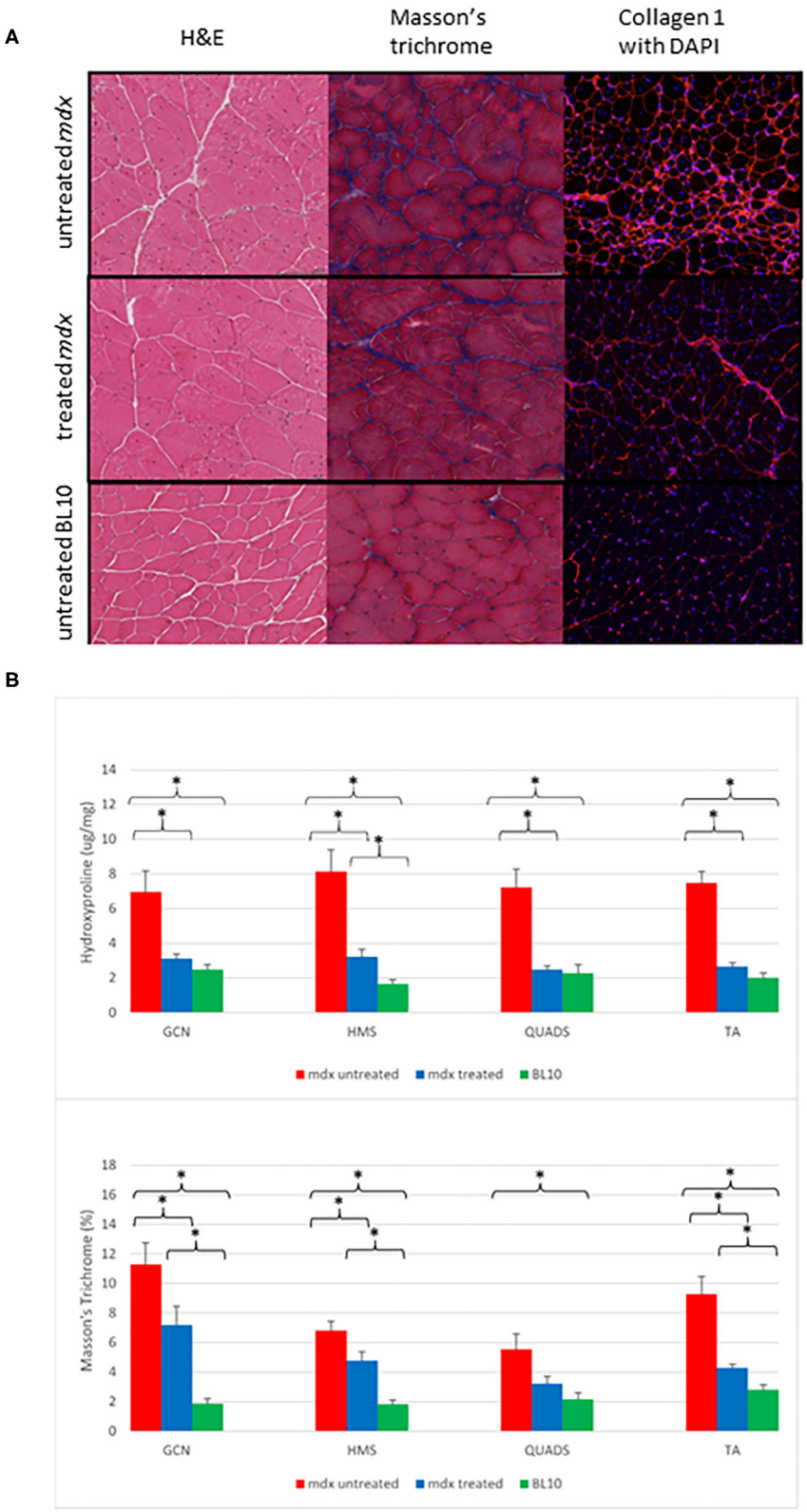


FIGURE 2 | (A) Examples of transverse sections from gastrocnemius (GCN) muscles. The left column shows H&E staining at ×20 magnification demonstrate that the *mdx* mouse groups have evidence of more centrally located nuclei, varying myocyte size and areas of fibrosis (white). The middle column demonstrates

(Continued)

FIGURE 2 | Masson's trichrome staining in ×20 magnification. Fibrosis is shown as dark blue, the mdx untreated group show diffuse fibrosis throughout with the BL10 group showing very small amounts of fibrosis, the treated mdx mouse group shows fewer areas of blue staining compared to the untreated mdx group. The right column is immunofluorescent staining for collagen 1 with DAPI, areas of red are areas with collagen 1, nuclei are highlighted as blue (Magnification ×10). Similar to the Masson's trichrome staining, the untreated *mdx* mice have large amounts of diffuse fibrosis evident across the whole slide, with the BL10 slide only showing collagen 1 staining in the sarcolemma in an ordered fashion. The treated *mdx* mice demonstrate an amount of collagen staining in between these. **(B)** Bar charts to show the differences in *ex vivo* quantification of fibrosis within muscle between treated *mdx*, untreated *mdx* and BL10 groups. **(A)** hydroxyproline quantification, **(B)** MaTr quantification. Significant differences are seen between the all muscles of the untreated *mdx* mice and the BL10 group using bother measures of fibrosis. Using Masson's trichrome (MaTr) quantification, all muscles had significantly lower amounts of fibrosis in the treated *mdx* group compared to the untreated. Hydroxyproline quantification identified 3/4 of these muscles as significantly lower in the treated mdx group. Error bars represent standard error. *$p < 0.05$, GCN, Gastrocnemius, HMS, Hamstrings, QUADS, Quadriceps, TA, Tibialis anterior.

as well as serum from a halofuginone naive *mdx* mouse. Quality control samples (3, 30, and 700 ng/mL) and mouse samples were mixed with organic solvent (methanol) containing a mixture of three generic internal standards. Samples were treated at a ratio of 3:1 solvent to sample. Following protein precipitation, samples were then centrifuged (30 min at 5,000 g). The resulting supernatant was diluted with double distilled water at a ratio of 2:1 parts of supernatant in a 96-well-plate. The plate was sealed, vortex mixed and analysed by LC-MS/MS with incurred samples quantified from the calibration line.

Functional Assessment of Grip Strength

Prior to the final scan, a measure of muscle function was assessed using a grip strength apparatus (BioSeb, Chaville, France), with both two and four limb assessment performed. These were tested using a "T" shaped bar attachment and were placed upon the apparatus before being pulled gently backwards by the tail. Mice had three attempts with at least a 1 min of rest between assessments. To reduce bias the same examiner performed the test, in the same temperature controlled, quiet environment. The weight and the maximal recorded value from the three attempts was used to calculate the normalised force (force/body weight) (25).

Statistical Analysis

All statistics were calculated using SPSS (version 23). Correlation is reported as a Pearson coefficient. Student's *t*-test was used to compare means, with $p < 0.05$ considered significant. One way anova was applied to compare grip test results with *post-hoc* analysis performed using the Tukey-Kramer test.

RESULTS

All mice (*mdx* $n = 21$, BL10 $n = 6$) tolerated halofuginone treatment, vehicle and EP3533 administration with no side effects attributed to the drug and no significant weight difference was seen between treated and untreated groups. There were no significant differences between BL10 mouse groups in terms of weight, functional assessment, *ex vivo* measures of fibrosis or baseline and follow up R1 change. The BL10 groups were therefore considered together for comparison with *mdx* mouse groups.

Halofuginone Treatment Effects

Pharmacokinetic results demonstrated that halofuginone was detectable in serum at six of the eight time points. The concentration range was from 2.85 to 21.2 ng/ml with the highest peak at 5 min post-administration, these findings suggest successful administration at a concentration expected to induce a treatment effect (**Figure 1**). Area under the curve analysis was 913.68 ng/ml/min.

Ex vivo Fibrosis Quantification

Histological analysis demonstrated the diffuse nature of the fibrotic changes in *mdx* mouse muscle and comparative differences between groups (**Figure 2**). All treated *mdx* muscles demonstrated significantly lower mean levels of fibrosis compared to untreated *mdx* using hydroxyproline assay. Comparing treated to untreated *mdx,* MaTr quantification was significantly lower in three of the four muscles, with the exception of the QUADS ($p = 0.06$). These results suggest that levels of fibrosis were lower in the majority of muscles in treated *mdx* groups compared to the untreated *mdx* groups (**Figure 3**).

Functional Assessment

The mouse groups were compared using a one-way Anova test. Using the two-limb grip strength measure the untreated *mdx* mouse group was significantly weaker than the other groups ($F = 6.08$, $p < 0.03$). *Post-hoc* analysis demonstrated a significant difference between the untreated and the treated *mdx* mice (untreated *mdx* 21.5 mN/g vs. treated *mdx* 37.0 mN/g). There were no statistically significant differences between the BL10 groups and the treated *mdx* mice ($p =$ NS).

R1 Change in EP3533 Enhanced MRI

Pre-treatment scans demonstrated no significant differences between the treated and untreated *mdx* mouse groups. **Figure 3A** shows representative T1 weighted axial MRI scans through hind limbs with regions of interest labelled. In the pre-treatment scans R1 change of all muscles were significantly higher at 60 min in the combined *mdx* groups than the BL10 at 60 min. These findings suggest that there was no detectable difference between the two *mdx* groups at baseline and that EP3533-enhanced scans were sensitive enough to discern between BL10 and *mdx* groups at an early age (12 weeks).

In the post-treatment scan, there were no significant differences between the untreated and treated BL10 groups in any of the muscles at any of the time points. No significant differences were seen between treated and untreated *mdx* groups in absolute R1 value at the post-treatment scan. When comparing the R1 change over time from pre-treatment to post-treatment scan there was a significant increase in three out of four of the muscle

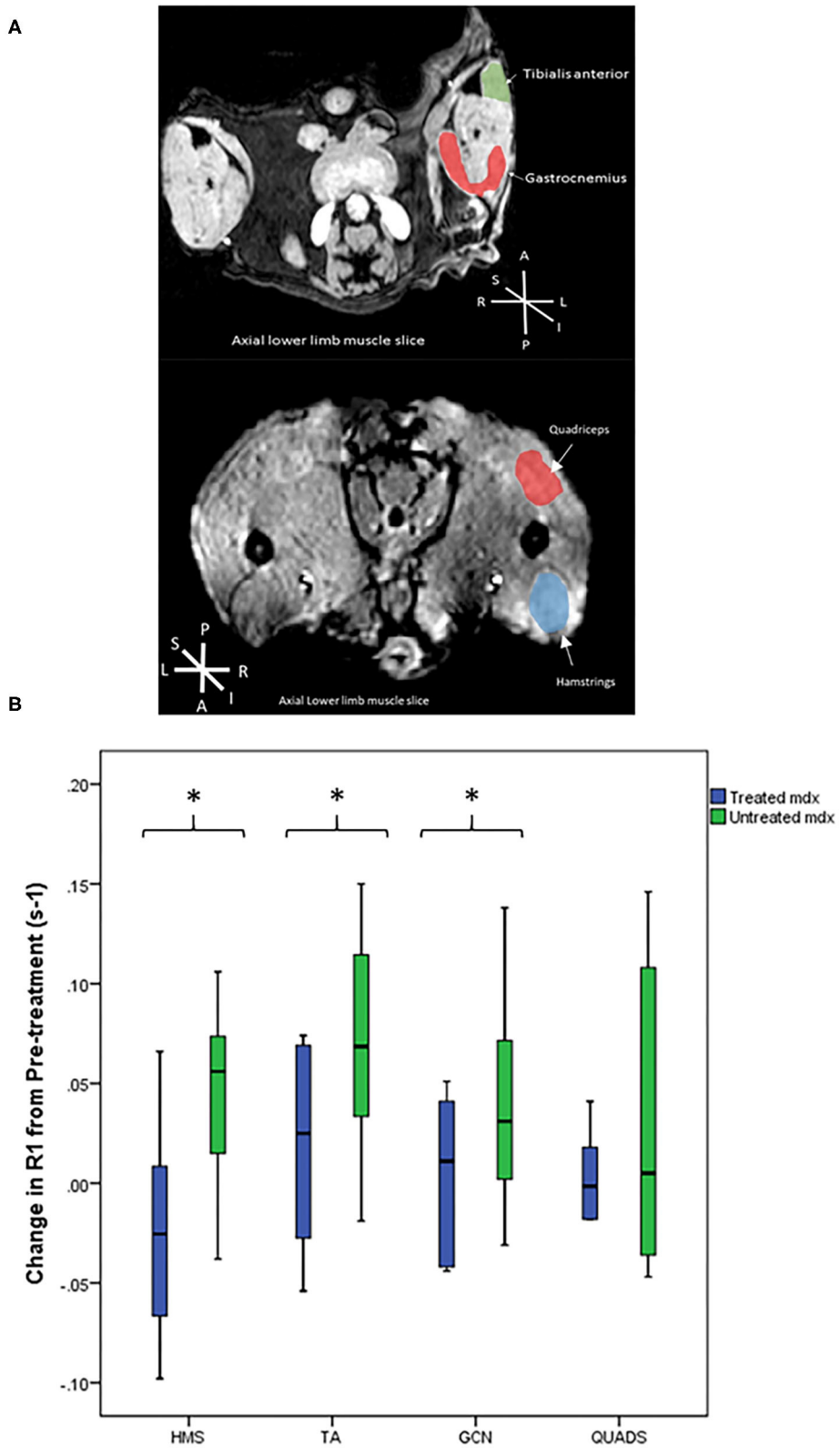


FIGURE 3 | (A) Representative T1 weighted axial MRI scans through hind limbs with regions of interest labelled. P, posterior, A, anterior, L, left, R, right, S, superior, I, inferior. **(B)** Box plots to compare the difference in R1 change between the *mdx* groups from pre-treatment to post-treatment scans. The treated *mdx* mice did not show significant increases in R1 post-treatment whereas the untreated *mdx* group showed significant increases in R1 change over time in all three muscles: gastrocnemius $p = 0.04$, hamstrings $p = 0.009$, and tibialis anterior $p = 0.01$. *$p < 0.05$, GCN, Gastrocnemius, HMS, Hamstrings, QUADS, Quadriceps, TA, Tibialis anterior.

of the untreated *mdx* group (GCN $p = 0.04$, HMS $p = 0.009$, and TA $p = 0.01$; **Table 1**, **Figure 3B**). No significant change was seen over time in any of the muscles in the other groups.

Correlation of EP3533-Enhanced MRI Findings to *ex vivo* Measures of Fibrosis

R1 values in all muscles at the post-treatment scan correlated significantly with Masson's trichrome (**Table 1**). R1 change at post-treatment correlated linearly in three out of four muscles with hydroxyproline. The strongest correlations to *ex vivo* measures of fibrosis were in the gastrocnemius muscle (**Table 1**).

Correlation of EP3533-Enhanced MRI to Functional Assessments

None of the two limb force assessments significantly correlated to the change in R1 values from the post-treatment scan, in either the individual muscles or all muscles grouped together. Considering all muscles together the change in R1 values including the pre and post-treatment scans correlated significantly, albeit weakly ($r = -0.33$, $p = 0.004$). The correlation may suggest that there is an association between R1 change at post treatment and muscle function, the lack of a strong correlation may be due to the other confounding factors that influence muscle function.

DISCUSSION

Muscular dystrophies such as DMD are slowly progressive conditions, the histological hallmarks include inflammation, fibrosis, and fat replacement of healthy muscle. Eventually, secondary to this pathology, the muscles are unable to function and clinical weakness is evident. Clinical trial outcome measures are biomarkers of disease progression that are increasingly important to test putative therapies. In clinical trials involving neuromuscular disease the most commonly used outcome measures are assessing muscle strength and muscle function. Use of a non-invasive outcome measure based on earlier pathology provides a potentially more sensitive biomarker to disease progression over a shorter period. Such outcome measures are sorely needed as muscular dystrophies may take several years to progress clinically, this makes drug development time consuming and expensive. This is the first study to look at use of EP3533 as an outcome measure in muscular dystrophy. Previous studies have shown that EP3533-enhanced MRI can demonstrate sufficient sensitivity to fibrosis *in vivo* to accurately stage disease, demonstrate disease progression and quantify fibrosis in organs with low proton density (9, 10, 12–14). This study showed that EP3533 was able to demonstrate a treatment effect as well as demonstrating strong correlations with *ex vivo* measures of fibrosis, and a weaker correlation to a functional assessment.

Halofuginone was chosen as an anti-fibrotic as it has previously demonstrated that it can induce a reduction in collagen and an improvement of muscle function in murine models (18, 19). Halofuginone is currently being evaluated in a clinical trial (26). This study further supported the use of halofuginone to reduce fibrosis *in vivo* and to improve muscle function. In this experiment the *mdx* treated and BL10 groups were not significantly different in a number of ways including: absolute values of R1 change over time, functional assessments, and degree of *ex vivo* fibrosis in several muscles (**Figure 3**). Other anti-fibrotics have been trialled in DMD patients, though in humans factors such as fat replacement may reduce the effectiveness of these drugs (27–29). Anti-fibrotics do have the advantage of not being mutation dependent, unlike a number of recent medications aimed at increasing dystrophin production (30–32). As anti-fibrotic agents are potential drug candidates to treat muscular dystrophies, we tested if EP3533-enhanced MRI could be used as an outcome measure, as it has been shown to be successful at monitoring fibrosis progression over time in murine models. A further advantage of EP3533 is that due to the non-invasive nature of the test fewer animals would need to be sacrificed to demonstrate interim changes in experiments.

TABLE 1 | Table to show absolute R1 change at both the pre- and post-treatment scans.

	Pre-treatment scan change in R1 from baseline (s^{-1})	Post-treatment scan change in R1 from baseline (s^{-1})	Correlation of change in R1 at post-treatment scan to Masson's Trichrome†	Correlation of change in R1 at post-treatment scan to hydroxyproline assay†
mdx TA	0.11 ± 0.05*	0.14 ± 0.06*	0.84**	0.58*
BL10 TA	0.11 ± 0.02	0.07 ± 0.01		
mdx GCN	0.10 ± 0.04*	0.10 ± 0.05*	0.80**	0.59*
BL10 GCN	0.04 ± 0.05	0.04 ± 0.05		
mdx HMS	0.11 ± 0.04*	0.12 ± 0.04*	0.59*	0.41*
BL10 HMS	0.03 ± 0.04	0.06 ± 0.06		
mdx QUADS	0.09 ± 0.12*	0.09 ± 0.05*	0.59*	0.24
BL10 QUADS	0.001 ± 0.12	0.04 ± 0.06		

Also demonstrates the high degree of significant Pearson correlation between R1 change at post-treatment scan and ex vivo measures of fibrosis.
TA, Tibialis anterior; GCN, Gastrocnemius; HMS, Hamstrings; QUADS, Quadriceps.
$P < 0.05$.
**$P < 0.001$.*
†*BL10 and mdx considered together.*

Currently the only validated way to quantify fibrosis in human skeletal muscle is via muscle biopsy (33). Muscle biopsy is invasive, risks sampling error and includes the risk of a general anaesthetic in patients with reduced respiratory function. EP3533, by contrast, has been shown to be able to measure change in skeletal muscle fibrosis over time in a relatively non-invasive manner. As seen in humans with muscular dystrophy, *mdx* mice demonstrate individual variation in fibrosis. This may explain why it was the *change* in R1 value from pre-treatment to post-treatment which demonstrated a significant difference between the treated and untreated groups rather than the absolute R1 change in the post-treatment scan alone (34, 35). This is important for future experiments using EP3533, with baseline and post treatment scans being the correct way to demonstrate change, rather than a single measurement.

Other quantitative MRI methods have been used clinically in muscular dystrophies, including native T1, T2, magnetic resonance spectroscopy, and Dixon fat fraction (FF) (36–41). In particular, FF has been shown to be a sensitive measure of disease progression. A drawback of the *mdx* mouse model is that it demonstrates minimal intramuscular fat replacement compared to humans with the condition. Therefore, the utility of EP3533-enhanced MRI scan would have to be compared against FF as an outcome measure. In human pathology, fat replacement eventually becomes the dominant pathology with large percentages of muscle replaced, therefore it is unknown whether EP3533-enhanced MRI is likely to be helpful in advanced disease or to be merely used alongside established biomarkers such as FF.

In murine experiments several standardised assessments have been developed to measure motor function and strength in a reliable way (25, 34, 42). External validity may be reduced due to different compensatory mechanisms employed by both species (43). Grip strength tests in rodents are most akin to quantitative muscle assessment in humans (44), with such grip assessments able to demonstrate significant differences between control and *mdx* mice at an early age (34). In this experiment functional results correlated significantly to R1 change ($r = -0.33$, $p = 0.004$), and demonstrated a significant difference between untreated and treated *mdx* groups ($p < 0.001$). In contrast to these results, Huebner et al. who performed halofuginone studies in *mdx* mice, did not report a significant difference between treated *mdx* mice and control in grip assessment (19). This difference may be due to the smaller group sizes used by Huebner et al. (19) ($n = 5$). As with functional outcome measures in humans, other factors such as cognition, stress and motivation can be confounding variables (35). In humans with DMD, quantitative MRI such as Dixon fat fraction (FF) measurement has been shown to have higher levels of correlation to the 6 min walk test (6MWT), ($r = -0.65$, $p < 0.001$), as well as other functional assessments (45–48). The correlations seen in this experiment were lower than that of FF in humans. Any future evaluation of efficacy of EP3533-induced R1 change in clinical trials should therefore be compared with functional assessments as well as MRI measures (45, 49). The relatively weak correlation between function and EP3533-induced R1 change may be due to the non-linear development of fibrosis throughout muscle, as seen in other human and murine studies (37, 50, 51).

There have been concerns around gadolinium-based contrast agents and their repeated use in humans due to potential neurotoxicity (52). Even less is known of the safety profile of peptide-based gadolinium agents which may be excreted less readily due to their complex structure. Further preclinical studies are required to provide safety data findings used to inform a future application for registration of potential human trials before this agent can be used as an outcome measure or in clinical practise.

CONCLUSION

As new therapies are trialled in muscular dystrophies there is great potential for EP3533 to be used as an outcome measure, adding important, quantifiable information relatively non-invasively. EP3533 can also be used to reduce animal numbers in murine experiments, allowing accurate quantification of fibrosis over time. Our study showed that EP3533-enhanced MRI can be used in a muscular dystrophy mouse model to demonstrate disease progression over time and to effectively monitor a treatment response.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The animal study was reviewed and approved by all procedures performed were in accordance with the ethical standards of directive 2010/63/EU of the European parliament and under the auspices of the terms of the animals (scientific procedures) act 1986 and project licence PB3CA650C, authorised by the Home Secretary, Home Office, United Kingdom.

AUTHOR CONTRIBUTIONS

All authors have contributed considerably to the design, execution of the study or to the analysis and interpretation of data, involved in drafting or critically revising the manuscript, and have read and approved its final version.

FUNDING

This work was supported by grants from Muscular Dystrophy UK (17GRO-PG12-0197) and from AFM-Téléthon (#20553) to VS.

REFERENCES

1. Bradley D, Parsons E. Newborn screening for duchenne muscular dystrophy. *Semi Neonatol.* (1998) 3:27–34. doi: 10.1016/S1084-2756(98)80146-2
2. Moat SJ, Bradley DM, Salmon R, Clarke A, Hartley L. Newborn bloodspot screening for duchenne muscular dystrophy: 21 years experience in Wales (UK). *Euro J Hum Genet.* (2013) 21:1049–53. doi: 10.1038/ejhg.2012.301
3. Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, et al. Diagnosis and management of duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurol.* (2010) 9:77–93. doi: 10.1016/S1474-4422(09)70271-6
4. Muntoni F, Torelli S, Ferlini A. Dystrophin and mutations: one gene, several proteins, multiple phenotypes. *Lancet Neurol.* (2003) 2:731–40. doi: 10.1016/S1474-4422(03)00585-4
5. Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, et al. Diagnosis and management of duchenne muscular dystrophy, part 2: implementation of multidisciplinary care. *Lancet Neurol.* (2010) 9:177–89. doi: 10.1016/S1474-4422(09)70272-8
6. Klingler W, Jurkat-Rott K, Lehmann-Horn F, Schleip R. The role of fibrosis in Duchenne muscular dystrophy. *Acta Myol.* (2012) 31:184–95.
7. Zanotti S, Gibertini S, Mora M. Altered production of extra-cellular matrix components by muscle-derived Duchenne muscular dystrophy fibroblasts before and after TGF-β1 treatment. *Cell Tissue Res.* (2010) 339:397–410. doi: 10.1007/s00441-009-0889-4
8. Desguerre I, Mayer M, Leturcq F, Barbet JP, Gherardi RK, Christov C. Endomysial fibrosis in duchenne muscular dystrophy: a marker of poor outcome associated with macrophage alternative activation. *J Neuropathol Exp Neurol.* (2009) 68:762–73. doi: 10.1097/NEN.0b013e3181aa31c2
9. Helm PA, Caravan P, French BA, Jacques V, Shen L, Xu Y, et al. Postinfarction myocardial scarring in mice: molecular MR imaging with use of a collagen-targeting contrast agent. *Radiology.* (2008) 247:788–96. doi: 10.1148/radiol.2473070975
10. Fuchs BC, Wang H, Yang Y, Wei L, Polasek M, Schühle DT, et al. Molecular MRI of collagen to diagnose and stage liver fibrosis. *J Hepatol.* (2013) 59:992–8. doi: 10.1016/j.jhep.2013.06.026
11. Polasek M, Fuchs BC, Uppal R, Schühle DT, Alford JK, Loving GS, et al. Molecular MR imaging of liver fibrosis: a feasibility study using rat and mouse models. *J Hepatol.* (2012) 57:549–55. doi: 10.1016/j.jhep.2012.04.035
12. Caravan P, Yang Y, Zachariah R, Schmitt A, Mino-Kenudson M, Chen HH, et al. Molecular magnetic resonance imaging of pulmonary fibrosis in mice. *Am J Respir Cell Mol Biol.* (2013) 49:1120–6. doi: 10.1165/rcmb.2013-0039OC
13. Polasek M, Yang Y, Schuhle DT, Yaseen MA, Kim YR, Sung YS, et al. Molecular MR imaging of fibrosis in a mouse model of pancreatic cancer. *Sci Rep.* (2017) 7:8114. doi: 10.1038/s41598-017-08838-6
14. Farrar CT, DePeralta DK, Day H, Rietz TA, Wei L, Lauwers GY, et al. 3D molecular MR imaging of liver fibrosis and response to rapamycin therapy in a bile duct ligation rat model. *J Hepatol.* (2015) 63:689–96. doi: 10.1016/j.jhep.2015.04.029
15. Murphy AP, Greally E, O'Hogain D, Blamire A, Caravan P, Straub V. Noninvasive quantification of fibrosis in skeletal and cardiac muscle in mdx mice using EP3533 enhanced magnetic resonance imaging. *Magn Reson Med.* (2019) 81:2728–35. doi: 10.1002/mrm.27578
16. McGaha TL, Phelps RG, Spiera H, Bona C. Halofuginone, an inhibitor of type-I collagen synthesis and skin sclerosis, blocks transforming-growth-factor-beta-mediated Smad3 activation in fibroblasts. *J Invest Dermatol.* (2002) 118:461–70. doi: 10.1046/j.0022-202x.2001.01690.x
17. Barzilai-Tutsch H, Bodanovsky A, Maimon H, Pines M, Halevy O. Halofuginone promotes satellite cell activation and survival in muscular dystrophies. *Biochim Biophys Acta.* (2016) 1862:1–11. doi: 10.1016/j.bbadis.2015.10.007
18. Turgeman T, Hagai Y, Huebner K, Jassal DS, Anderson JE, Genin O, et al. Prevention of muscle fibrosis and improvement in muscle performance in the mdx mouse by halofuginone. *Neuromuscul Disord.* (2008) 18:857–68. doi: 10.1016/j.nmd.2008.06.386
19. Huebner KD, Jassal DS, Halevy O, Pines M, Anderson JE. Functional resolution of fibrosis in mdx mouse dystrophic heart and skeletal muscle by halofuginone. *Am J Physiol Heart Circ Physiol.* (2008) 294:H1550–61. doi: 10.1152/ajpheart.01253.2007
20. Sigma-Aldrich. *Halofuginone Bromide.* (2018). Available online at: https://www.sigmaaldrich.com/catalog/search?term=Halofuginone&interface=All&N=0&mode=partialmax&lang=en&region=GB&focus=product (accessed May 8, 2018).
21. Niskanen J. *Aedes a Tool for Analyzing Medical Images.* (2016). Available online at: http://aedes.uef.fi/ (accessed February 1, 2016).
22. Zhang J, Zhang G, Morrison B, Mori S, Sheikh KA. Magnetic resonance imaging of mouse skeletal muscle to measure denervation atrophy. *Exp Neurol.* (2008) 212:448–57. doi: 10.1016/j.expneurol.2008.04.033
23. Carlson G. *Determination of Hydroxyproline Content as a Measure of Fibrosis in Nondystrophic and Dystrophic Skeletal Muscle.* (2014). Available online at: http://www.treat-nmd.eu/downloads/file/sops/dmd/MDX/DMD_M.1.2.006.pdf (accessed May 31, 2016).
24. UK Home Office. *Mouse: Decision Tree for Blood Sampling.* (2018). Available online at: https://www.nc3rs.org.uk/mouse-decision-tree-blood-sampling (accessed May 9, 2018).
25. De Luca A, Tinsley J, Aartsma-Rus A, van Putten M, Nagaraju K, de La Porte S, et al. *Use of Grip Strength Meter to Assess the Limb Strength of mdx Mice.* (2014). Available online at: http://www.treatnmd.eu/downloads/file/sops/dmd/MDX/DMD_M.2.2.001.pdf (accessed March 13, 2018).
26. Clinical Trials.gov. *Safety, Tolerability, and Pharmacokinetics of Single and Multiple Doses of HT-100 in Duchenne Muscular Dystrophy.* (2018). Available online at: https://clinicaltrials.gov/ct2/show/NCT01847573 (accessed January 9, 2019).
27. Nio Y, Tanaka M, Hirozane Y, Muraki Y,Okawara M, Hazama M, Matsuo T. Phosphodiesterase 4 inhibitor and phosphodiesterase 5 inhibitor combination therapy has anti-fibrotic and anti-inflammatory effects in mdx mice with duchenne muscular dystrophy. *FASEB J.* (2017) 31:5307–20. doi: 10.1096/fj.201700249R
28. Zanotti S, Mora M. Test of anti-fibrotic drugs in a cellular model of fibrosis based on muscle-derived fibroblasts from duchenne muscular dystrophy patients. *Methods Mol Biol.* (2018) 1687:205–17. doi: 10.1007/978-1-4939-7374-3_15
29. Swiderski K, Todorov M, Gehrig SM, Naim T, Chee A, Stapleton DI et al. Tranilast administration reduces fibrosis and improves fatigue resistance in muscles of mdx dystrophic mice. *Fibrogenesis Tissue Repair.* (2014) 7:1. doi: 10.1186/1755-1536-7-1
30. van Deutekom JCT, Bremmer-Bout M, Janson AAM, Ginjaar LB, Baas F, den Dunnen JT, et al. Antisense-induced exon skipping restores dystrophin expression in DMD patient derived muscle cells. *Hum Mol Genet.* (2001) 10:1547–54. doi: 10.1093/hmg/10.15.1547
31. McDonald CM, Campbell C, Torricelli RE, Finkel RS, Flanigan KM, Goemans N, et al. Ataluren in patients with nonsense mutation duchenne muscular dystrophy (ACT DMD): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet.* (2017) 390:1489–98. doi: 10.1016/S0140-6736(17)31611-2
32. Voit T, Topaloglu H, Straub V, Mutoni F, Deconinck N, Campion G, et al. Safety and efficacy of drisapersen for the treatment of duchenne muscular dystrophy (DEMAND II): an exploratory, randomised, placebo-controlled phase 2 study. *Lancet Neurol.* (2014) 13:987–96. doi: 10.1016/S1474-4422(14)70195-4
33. Sardone V, Ellis M, Torelli S, Feng L, Chambers D, Eastwood D et al. A novel high-throughput immunofluorescence analysis method for quantifying dystrophin intensity in entire transverse sections of duchenne muscular dystrophy muscle biopsy samples. *PLoS ONE.* (2018) 13:e0194540. doi: 10.1371/journal.pone.0194540
34. Spurney CF, Gordish-Dressman H, Guerron AD, Sali A, Pandey GS, Rawat R, et al. Preclinical drug trials in the mdx mouse: assessment of reliable and sensitive outcome measures. *Muscle Nerve.* (2009) 39:591–602. doi: 10.1002/mus.21211
35. Grounds MD, Radley HG, Lynch GS, Nagaraju K, De Luca A. Towards developing standard operating procedures for pre-clinical testing in the mdx mouse model of duchenne muscular dystrophy. *Neurobiol Dis.* (2008) 31:1–19. doi: 10.1016/j.nbd.2008.03.008
36. Barnard AM, Willcocks RJ, Finanger EL, Daniels MJ, Triplett WT, Rooney WD, et al. Skeletal muscle magnetic resonance biomarkers correlate with function and sentinel events in duchenne muscular dystrophy. *PLoS ONE.* (2018) 13:e0194283. doi: 10.1371/journal.pone.0194283

37. Arpan I, Forbes SC, Lott DJ, Senesac CR, Daniels MJ, Triplett WT et al. T2 mapping provides multiple approaches for the characterization of muscle involvement in neuromuscular diseases: a cross-sectional study of lower leg muscles in 5–15-year-old boys with duchenne muscular dystrophy. *NMR Biomed.* (2012) 26:320–28. doi: 10.1002/nbm.2851
38. Carlier PG, Marty B, Scheidegger O, de Sousae PL, Baudinf PY, Snezhkoc E, et al. Skeletal muscle quantitative nuclear magnetic resonance imaging and spectroscopy as an outcome measure for clinical trials. *J Neuromuscul Dis.* (2016) 3:1–28. doi: 10.3233/JND-160145
39. Fatehi F, Salort-Campana E, Le Troter A, Lareau-Trudel E, Bydder M, Foure' A, et al. Long-term follow-up of MRI changes in thigh muscles of patients with facioscapulohumeral dystrophy: a quantitative study. *PLoS ONE.* (2017) 12:e0183825. doi: 10.1371/journal.pone.0183825
40. Murphy AP, Morrow J, Dahlqvist JR, Stojkovic T, Willis TA, Sinclair CDJ, et al. Natural history of limb girdle muscular dystrophy R9 over 6 years: searching for trial endpoints. *Ann Clin Transl Neurol.* (2019) 6:1033–45. doi: 10.1002/acn3.774
41. Yin L, Xie Z-y, Xu H-y, Zheng SS, Wang ZX, Xiao JX, et al. T2 mapping and fat quantification of thigh muscles in children with duchenne muscular dystrophy. *Curr Med Sci.* (2019) 39:138–45. doi: 10.1007/s11596-019-2012-8
42. Treat NMD. *TREAT NMD: Experimental Protocols for DMD Animal Models.* (2016). Available online at: http://www.treat-nmd.eu/research/preclinical/dmd-sops/ (accessed January 8, 2019).
43. Hu X, Charles JP, Akay T, Hutchinson JR, Blemker SS. Are mice good models for human neuromuscular disease? Comparing muscle excursions in walking between mice and humans. *Skeletal Muscle.* (2017) 7:26. doi: 10.1186/s13395-017-0143-9
44. McDonald CM, Henricson EK, Abresch RT, Florence JM, Eagle M, Gappmaier E, et al. The 6 minute walk test and other endpoints in Duchenne muscular dystrophy: Longitudinal natural history observations over 48 weeks from a multicenter study. *Muscle Nerve.* (2013) 48:343–56. doi: 10.1002/mus.23902
45. Fischer D, Hafner P, Rubino D, Schmid M, Neuhaus C, Jung H, et al. The 6-minute walk test, motor function measure and quantitative thigh muscle MRI in Becker muscular dystrophy: a cross-sectional study. *Neuromusc Disord.* (2016) 26:414–22. doi: 10.1016/j.nmd.2016.04.009
46. Wokke BH, van den Bergen JC, Versluis MJ, Niks EH, Milles J, Webb AG, et al. Quantitative MRI and strength measurements in the assessment of muscle quality in Duchenne muscular dystrophy. *Neuromusc Disord.* (2014) 24:409–16. doi: 10.1016/j.nmd.2014.01.015
47. Willcocks RJ, Rooney WD, Triplett WT, Forbes SC, Lott DJ, Senesac CR, et al. Multicenter prospective longitudinal study of magnetic resonance biomarkers in a large duchenne muscular dystrophy cohort. *Ann Neurol.* (2016) 79:535–47. doi: 10.1002/ana.24599
48. Willcocks RJ, Triplett WT, Forbes SC, Arora H, Senesac CR, Lott DJ et al. Magnetic resonance imaging of the proximal upper extremity musculature in boys with duchenne muscular dystrophy. *J Neurol.* (2017) 264:64–71. doi: 10.1007/s00415-016-8311-0
49. Akima H, Lott D, Senesac C, Deol J, Germain S, Arpan I, et al. Relationships of thigh muscle contractile and non-contractile tissue with function, strength, and age in boys with duchenne muscular dystrophy. *Neuromusc Disord.* (2012) 22:16–25. doi: 10.1016/j.nmd.2011.06.750
50. Vohra R, Batra A, Forbes SC, Vandenborne K, Walter GA. Magnetic resonance monitoring of disease progression in *mdx* mice on different genetic backgrounds. *Am J Pathol.* (2017) 187:2060–70. doi: 10.1016/j.ajpath.2017.05.010
51. Gutpell KM, Hrinivich WT, Hoffman LM. Skeletal muscle fibrosis in the mdx/utrn+/− mouse validates its suitability as a murine model of duchenne muscular dystrophy. *PLoS ONE.* (2015) 10:e0117306. doi: 10.1371/journal.pone.0117306
52. Dekkers IA, Roos R, van der Molen AJ. Gadolinium retention after administration of contrast agents based on linear chelators and the recommendations of the European medicines agency. *Euro Radiol.* (2018) 28:1579–84. doi: 10.1007/s00330-017-5065-8

Conflict of Interest: PC has equity in, and is a consultant to Collagen Medical LLC which holds the patent rights to EP-3533.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neurology

ORIGINAL RESEARCH
published: 03 September 2021
doi: 10.3389/fneur.2021.656487

Assessing Fatty Infiltration of Paraspinal Muscles in Patients With Lumbar Spinal Stenosis: Goutallier Classification and Quantitative MRI Measurements

Filippo Mandelli[1,2,3,4]*, *Corina Nüesch*[1,3,4,5], *Yuancheng Zhang*[1], *Florian Halbeisen*[6], *Stefan Schären*[1,3,4], *Annegret Mündermann*[1,3,4,5] and *Cordula Netzer*[1,3,4]

[1] Department of Spine Surgery, University Hospital Basel, Basel, Switzerland, [2] Department of Orthopaedics, University Children's Hospital Basel, Basel, Switzerland, [3] Department of Biomedical Engineering, University of Basel, Basel, Switzerland, [4] Department of Clinical Research, University of Basel, Basel, Switzerland, [5] Department of Orthopaedics and Traumatology, University Hospital Basel, Basel, Switzerland, [6] Institute for Clinical Epidemiology and Biostatistics, University Hospital Basel, Basel, Switzerland

OPEN ACCESS

Edited by:
Jordi Diaz-Manera,
University of Newcastle,
United Kingdom

Reviewed by:
Anna Pichiecchio,
Neurological Institute Foundation Casimiro Mondino (IRCCS), Italy
Bahar Shahidi,
University of California, San Diego, United States
Rafael Menezes-Reis,
Federal University of Amazonas, Brazil

***Correspondence:**
Filippo Mandelli
filippo.mandelli@usb.ch

Specialty section:
This article was submitted to Neuromuscular Disorders and Peripheral Neuropathies, a section of the journal Frontiers in Neurology

Received: 20 January 2021
Accepted: 23 July 2021
Published: 03 September 2021

Citation:
Mandelli F, Nüesch C, Zhang Y, Halbeisen F, Schären S, Mündermann A and Netzer C (2021) Assessing Fatty Infiltration of Paraspinal Muscles in Patients With Lumbar Spinal Stenosis: Goutallier Classification and Quantitative MRI Measurements. Front. Neurol. 12:656487. doi: 10.3389/fneur.2021.656487

Objective: Fatty infiltration of paraspinal muscle is associated with spinal disorders. It can be assessed qualitatively (i.e., Goutallier classification) and quantitatively using image processing software. The aims of this study were to compare paraspinal muscle fatty infiltration as assessed using the Goutallier classification vs. quantitative magnetic resonance images (MRI) measurements and to investigate the association between anthropometric parameters and paraspinal muscle morphology and fatty infiltration in patients with symptomatic lumbar spinal stenosis (LSS).

Methods: Patients affected by symptomatic LSS scheduled for surgery with available MRI of the lumbar spine were included in this retrospective cross-sectional study. Fatty infiltration at each lumbar level was rated qualitatively according to the Goutallier classification and quantified based on the cross-sectional area (CSA) of the paraspinal muscle, of its lean fraction (LeanCSA), and the ratio between LeanCSA and CSA and the CSA relative to the CSA of vertebral body (RCSA). Considering the muscle as a single unit, overall fatty infiltration according to Goutallier, overall CSA, LeanCSA, LeanCSA/CSA, and RCSA were computed as averages (aGoutallier, aCSA, aLeanCSA, aLeanCSA/aCSA, and aRCSA). Associations among parameters were assessed using Spearman's respective Pearson's correlation coefficients.

Results: Eighteen patients, with a mean age of 71.3 years, were included. aGoutallier correlated strongly with aLeanCSA and aLeanCSA/aCSA ($R = -0.673$ and $R = -0.754$, both $P < 0.001$). There was a very strong correlation between values of the left and right sides for CSA ($R = 0.956$, $P < 0.001$), LeanCSA ($R = 0.900$, $P < 0.001$), and LeanCSA/CSA ($R = 0.827$, $P < 0.001$) at all levels. Among all anthropometric measurements, paraspinal muscle CSA correlated the most with height (left: $R = 0.737$, $P < 0.001$; right: $R = 0.700$, $P < 0.001$), while there was a moderate correlation between vertebral body CSA and paraspinal muscle CSA (left: $R = 0.448$, $P < 0.001$;

right: $R = 0.454$, $P < 0.001$). Paraspinal muscle CSA correlated moderately with body mass index (BMI; left: $R = 0.423$, $P < 0.001$; right: $R = 0.436$, $P < 0.001$), and there was no significant correlation between aLeanCSA or aLeanCSA/CSA and BMI.

Conclusions: The Goutallier classification is a reliable yet efficient tool for assessing fatty infiltration of paraspinal muscles in patients with symptomatic LSS. We suggest taking body height as a reference for normalization in future studies assessing paraspinal muscle atrophy and fatty infiltration.

Keywords: paraspinal muscles fatty infiltration, muscle fatty infiltration, Goutallier, paraspinal cross-sectional area, paraspinal muscles morphology

INTRODUCTION

Muscle impairment is an important component of spinal disorders. In fact, muscle dysfunction can be considered both a cause and a consequence of low back pain (LBP) (1)—for instance, several studies have shown an association between fatty infiltration and the reduced cross-sectional area (CSA) of paraspinal muscles with chronic LBP (2–7). Moreover, fatty degeneration may lead to functional limitations such as deterioration of balance and alignment of the spine (8, 9). However, to date, the contribution of compromised muscles to pain in spinal pathologies, such as facet arthropathy, disc degeneration, spinal stenosis, or deformity, is unclear. Paraspinal muscle fatty degeneration and atrophy have been reported in patients with LBP and disc herniation (10–13) or degenerative lumbar flat back (14). In patients with disc herniation, muscle infiltration may be asymmetric and more pronounced on one side than the other (15). Moreover, greater fatty infiltration in muscles such as multifidus, longissimus, and psoas correlates with poorer functional outcomes with an even stronger relation in persons with a history of LBP (16). Similarly, patients with lumbar spinal stenosis (LSS) or symptomatic LSS (sLSS) show increased muscle atrophy and fatty infiltration (17). LSS is one of the most frequent spinal disorders and the most common reason for spine surgery in the population aged above 65 years (18). While there are several studies investigating the correlation of muscle composition and morphology in patients with LBP, to date, evidence for this relationship in patients with LSS is lacking.

Fatty degeneration of the multifidus and reduced CSA of the psoas muscles have been associated with a lower functional performance in terms of higher scores on the Oswestry Disability Index (ODI) (19–21). In addition, a reduced CSA of the multifidus may predispose patients to worse outcome after surgery for sLSS (22). Moreover, the relation between CSA of paraspinal muscles and anthropometric parameters has been explored. The results of previous studies are inconclusive, where some studies did not observe a significant association between paraspinal muscle CSA and body height or body mass (23–25), while others reported a greater paraspinal CSA in taller and heavier persons (26).

Magnetic resonance imaging (MRI) facilitates the detailed investigation of the lumbar spine and the estimation of the morphology and composition of paraspinal muscles, including multifidus and erector spinae (longissimus and iliocostalis). These assessments can be performed in a qualitative or in a quantitative way. The Goutallier classification is a visual grading system to qualitatively assess fatty infiltration (27–30). Initially proposed for grading fatty degeneration of the rotator cuff muscles on computer tomography, the Goutallier classification has been expanded to MRI and to the evaluation of other muscles, including back muscles (6, 31, 32). Moreover, there is evidence of positive correlations between Goutallier grades and the severity of disc degeneration as well as age (33). Previous studies have shown a substantial to excellent intraobserver and a good interobserver reliability (32, 33) as well as a significant positive correlation of Goutallier grades, with the percentage of fat in the multifidus muscle measured in a quantitative way (32). Although the Goutallier classification is simple to apply and a useful tool for clinicians when evaluating their patients, it has the disadvantages of qualitative measurements and ordinal scales. This classification depends on the experience of the assessor and is reported on an ordinal scale with five discrete levels. In contrast, the quantitative MRI measurements of paraspinal muscles overcome such limitations by being objective and continuous measures. However, these quantitative measures are more time consuming and hence more suitable for research rather than clinical settings. To date, data on comparisons between qualitative and quantitative assessments of paraspinal muscle fatty infiltration and their association with patient characteristics in patients with sLSS are lacking. The aims of this study were to compare paraspinal muscle fatty infiltration as assessed using the Goutallier classification vs. quantitative MRI measurements, assess asymmetry in muscle degeneration, and investigate the association between anthropometric parameters and paraspinal muscle morphology and fatty infiltration in patients with sLSS.

MATERIALS AND METHODS

This single center cross-sectional study was approved by the regional ethics committee and conducted in accordance with the Declaration of Helsinki. All enrolled patients provided written informed consent.

Study Cohort

Patients with a diagnosis of sLSS scheduled for decompressive surgery at the University Hospital Basel from April 2019 to

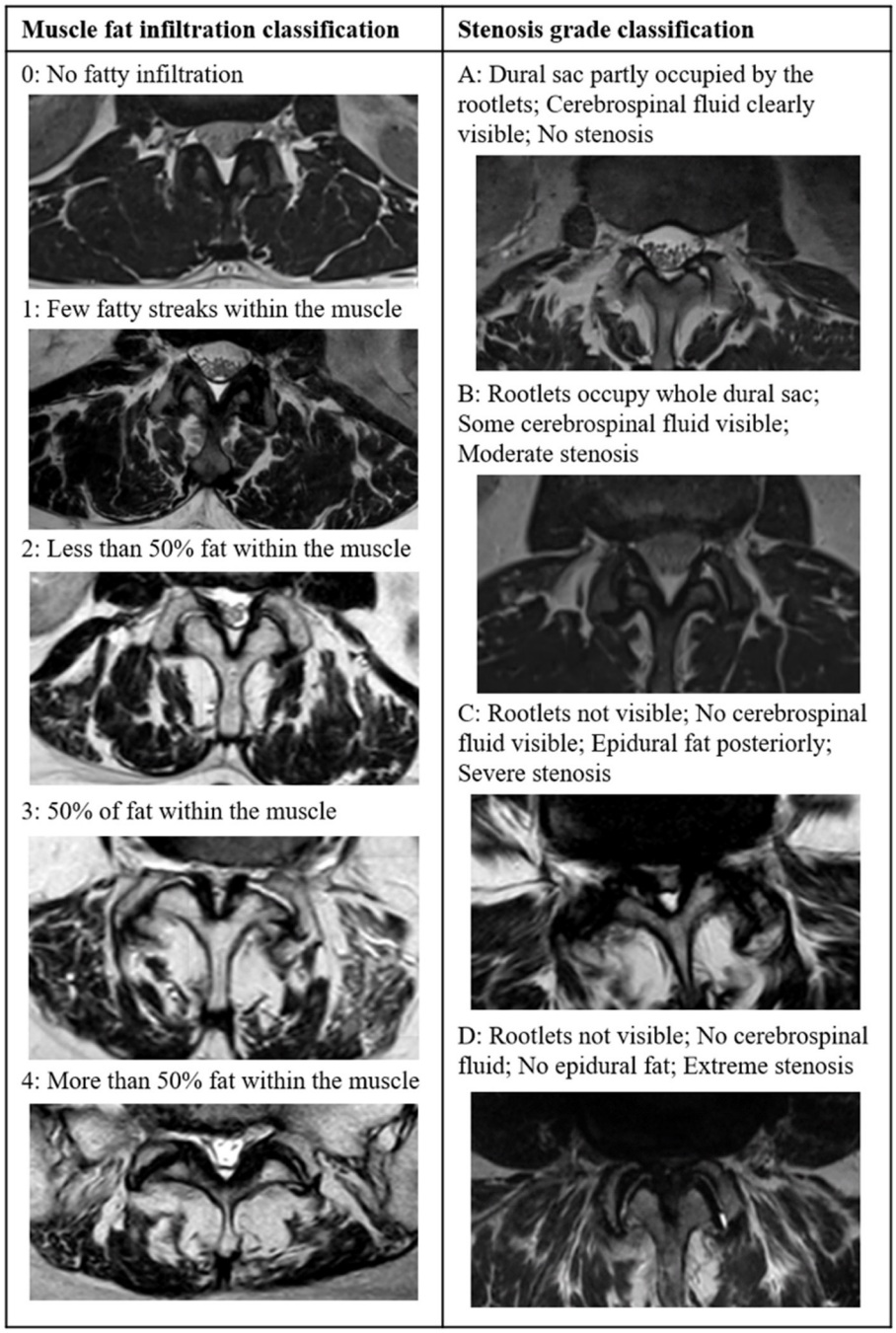


FIGURE 1 | T2-weighted axial images showing the grading of fatty infiltration of paraspinal muscles according to the Goutallier classification (left) (27) and grading of lumbar spinal stenosis according to the Schizas classifications (right) (36).

August 2020 were screened for this study. The participants were recruited in the context of a larger clinical research project, and eligible patients were informed about the study after admission to the hospital on the preoperative day. The inclusion criteria were as follows: diagnosed sLSS and availability of MR images of the lumbar spine from L1 to S1. The exclusion criteria were the following: prior surgery of the lumbar spine, additional pathologies that influence the mobility of the pelvis (such as internal fixation of the sacro-iliac joint or hip disorders affecting the gait), use of walking aids, and inability to provide informed consent. Age, sex, body mass, body height, and body mass index (BMI) were recorded. The level of LBP of the participants and the extent to which the pain impacts their daily activities and social life were estimated using the validated German version of

the standardized questionnaire ODI (34, 35). The ODI comprises 10 self-administered items describing the pain and limitations experienced when performing daily activities: pain intensity, personal care, lifting, walking, sitting, standing, sleeping, sexual life, social life, and traveling. The resulting score ranges from 0 (no impact) to 100 (bed-bound and extremely limited).

Lumbar MRI

All patients received MRI of the lumbar spine for clinical purposes. All MRI examinations included at least a sagittal T1- or T2-weighted sequence that was used to define the corresponding axial cut to be measured and an axial T2-weighted sequence to perform the qualitative and quantitative measurement of muscle morphology and composition. The MR images were obtained at our clinic (Prisma 3T, Siemens Healthineers, Erlangen, Germany) or provided by external providers at first consultation.

Qualitative Assessment of Paraspinal Muscle Fatty Infiltration

A qualitative assessment of paraspinal muscle fatty infiltration was performed using the Goutallier classification system (27). Accordingly, the muscle composition of the paraspinal muscles multifidus and longissimus on MRI was classified independently by two readers (FM and YZ) into five different grades based on the visually assessed fat/muscle ratio at each disc level from L1/L2 to L5/S1 (five segments in total; **Figure 1**). The inter-reader reliability was 0.701 (Cohens Kappa, $P < 0.001$). In case of disagreement between assessments, consensus was reached by a third reader (CNe). The grades range from grade 0—no visible fatty infiltration to grade 4—more than 50% of fat within the muscle. Overall qualitative fatty infiltration was computed as average Goutallier (aGoutallier) of all segments because fatty infiltration measured at each segment presumably affects the function of the entire muscle.

Quantitative Assessment of Paraspinal Muscle Fatty Infiltration

The quantitative assessment of paraspinal muscle fatty infiltration was performed using ImageJ image analysis software (version 1.52t, National Institutes of Health, Bethesda, Maryland) according to Fortin et al. (20, 37). ImageJ has been used in previous studies to assess fatty infiltration in MRI series of the lumbar spine (20, 32, 38, 39) and has shown a good intra- and interobserver reliability (37). Measurements of paraspinal muscle fatty infiltration were performed independently by two assessors (FM and YZ). A T2-weighted axial image was selected at each vertebral body level in the center of the body itself as identified on the sagittal-view image. For L1 to L5, we measured the CSA of the paraspinal muscles on each side, including the multifidus and the erector spinae (longissimus and iliocostalis) muscles, and the CSA of the vertebral body. The inter-reader reliability was very high [intraclass correlation coefficient, ICC (95% confidence interval)]—CSA: 0.939 (0.917, 0.955) and vertebral body CSA: 0.894 (0.722, 0.944). The relative CSA (RCSA) was defined as the ratio between muscle CSA and vertebral body CSA and calculated for each level and side. Thresholding of grayscale for lean muscle was repeated for each level, and each image was segmented according to the specific threshold (**Figure 2**). The CSA of lean muscle in the region of interest was defined as LeanCSA and measured on each side. The inter-reader reliability of LeanCSA was ICC (95% confidence interval)—LeanCSA: 0.959 (0.944, 0.970). The ratio of LeanCSA to the paraspinal muscle CSA was defined as functional CSA (LeanCSA/CSA), represented as percent of muscle CSA and calculated for each level and side. The overall CSA, RCSA, and LeanCSA were computed as average CSA (aCSA), average RCSA (aRCSA), and average LeanCSA (aLeanCSA) across all levels, considering the muscle as a single unit for each side. The total muscle CSA, total RCSA, and total LeanCSA at each level were calculated as the sum of the left and right muscle CSA, RCSA, and LeanCSA, respectively. Overall LeanCSA/CSA was computed as average LeanCSA/CSA (aLeanCSA/aCSA) of all segments and both sides.

Radiological Assessment of Severity of Spinal Stenosis

We graded the severity of the spinal stenosis according to the Schizas classification (36). The grading system is based on cerebral spine fluid/rootlet ratio on axial T2 images and consists of seven grades (A1, A2, A3, A4, B, C, and D) of stenosis (**Figure 2**). The severity of stenosis was assessed for each lumbar level addressed during the scheduled surgery and carried out by one reader (FM). The number of levels with stenosis and the highest severity of stenosis were determined and used in the analysis. A1 to A4 grades were grouped into a single group A.

Statistical Analysis

Statistical analyses were performed in SPSS Statistics, version 27 (IBM Corporation, Armonk, New York, USA). All data were checked for normality using the Kolmogorov–Smirnov test. The descriptive statistics for normally distributed parameters were performed using mean and one standard deviation (SD) and for all others as median and interquartile range (IQR). Differences in CSA, RCSA, LeanCSA, and LeanCSA/CSA between levels were identified for each side using analysis of variance (ANOVA) for repeated measures, with levels as within-subject factor, and upon significant results, *t*-tests for dependent samples were performed as *post-hoc* tests for comparisons between pairs of levels. The associations between continuous and normally distributed parameters were detected using Pearson's correlation coefficient. The associations between or with ordinal (Goutallier grade) and/or not normally distributed parameters were detected using Spearman's correlation coefficient. All correlations were performed separately for each side, except for correlations between parameters describing average muscle atrophy and fatty infiltration where the combined values for both sides were included. Correlations were considered very weak for $0 \leq |R| < 0.2$, weak for $0.2 \leq |R| < 0.4$, moderate for $0.4 \leq |R| < 0.6$, strong for $0.6 \leq |R| < 0.8$, and very strong for $0.8 \leq |R| \leq 1.0$ (40). The significance level for all tests was set *a priori* to 0.05.

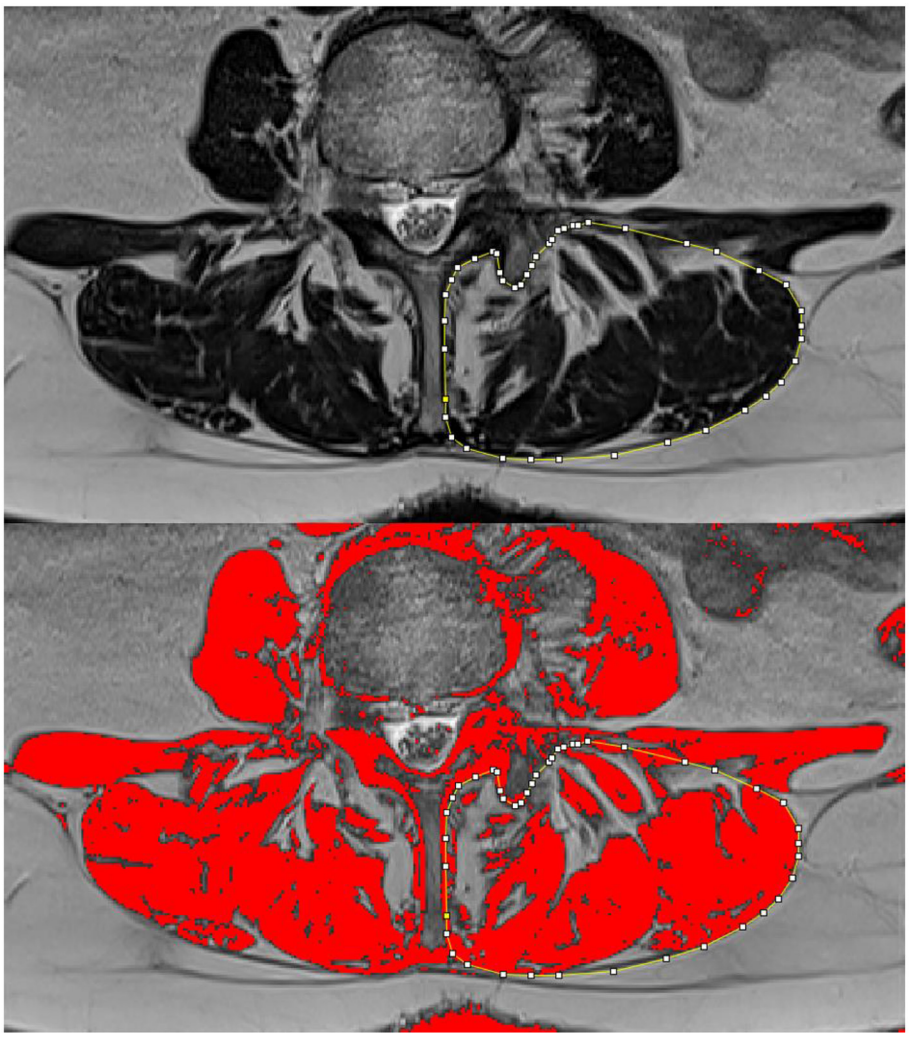


FIGURE 2 | Using the ImageJ analysis software, the left paraspinal muscles are contoured and the cross-sectional area is measured (top); after thresholding, the lean muscle is represented in red, and its area (LeanCSA) is measured (bottom).

RESULTS

Patient Characteristics

Eighteen patients with a mean age of 71.3 years (SD: 8.4) were included. The proportion of male patients was 45% (eight of 18 patients). The mean body mass was 75.8 kg (SD: 15.0). The mean height was 167.1 cm (SD: 8.6). The mean BMI of all patients was 27.0 kg/m^2 (SD: 3.9). The mean ODI score was 28.7 (SD: 13.5). The highest severity of LSS across all levels was grade B in four (22.2%) patients, grade C in 12 (66.7%) patients, and grade D in two (11.1%) patients. None of the patients had the highest stenosis severity grade A.

Qualitative Assessment of Paraspinal Muscle Fatty Infiltration

The median fatty infiltration according to the Goutallier classification system among all levels in all patients was 2.0 (IQR: 1.0–3.0). **Figure 3** shows the Goutallier grade at each level. The mean aGoutallier of all patients was 1.7 (SD: 0.6).

Quantitative Assessment of Paraspinal Muscles

The MR images of two subjects did not include an axial cut of the L1 level; data at this level for 16 of the 18 patients were included. The sum of the left and right total aCSA of the paraspinal muscles was 54.2 cm^2 (SD: 9.3), with the highest total paraspinal muscle CSA value of 58.0 cm^2 (SD: 9.9) at the L3 level and the lowest value of 45.4 cm^2 (SD: 20.0) at the L1 level (**Figure 4**). The statistically significant differences in paraspinal muscle CSA between levels are indicated in **Figure 4**. The total aRCSA was 4.4 (SD: 0.7), with the highest total (sum of left and right) RCSA value of 4.5 (SD: 0.7) at L2 level, although the differences between levels were not statistically significant (**Figure 4**). The total aLeanCSA was 27.7 cm^2 (SD: 10.0), with the highest total (sum of left and

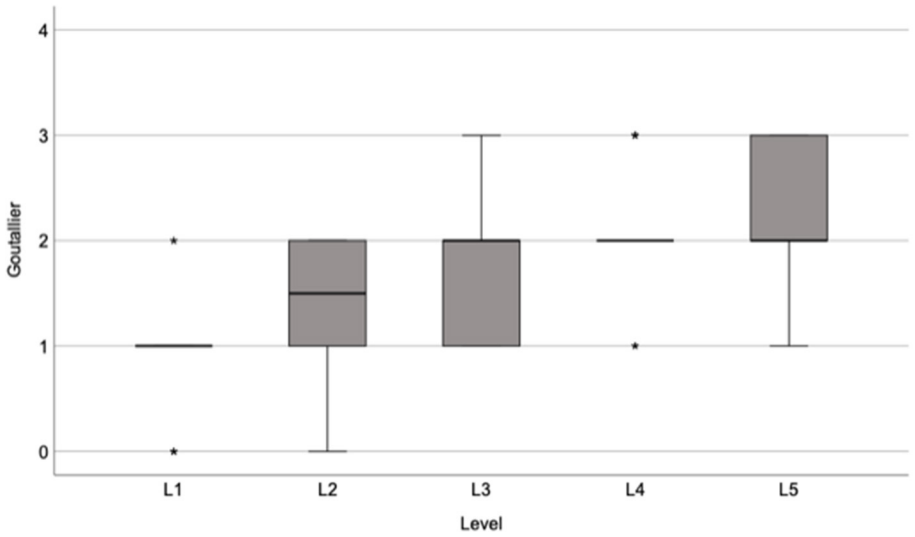


FIGURE 3 | Box plot showing the median and interquartile range of Goutallier grades at each level of the lumbar spine.

right) LeanCSA value of 30.9 cm^2 (SD: 10.9) at L2 and the lowest value of 22.8 cm^2 (SD: 8.2) at L5 (**Figure 4**). The statistically significant differences in LeanCSA between levels are indicated in **Figure 4**. The total aLeanCSA/aCSA was 50.2% (SD: 12.0%), where the values decreased from L1 to L5, with the highest value of 57.8% (SD: 12.6%) at L1 and the lowest value of 42.2% (SD: 13.3%) at L5 (**Figure 4**). The statistically significant differences in LeanCSA/CSA between levels are indicated in **Figure 4**.

Association Between Qualitative and Quantitative Measures of Fatty Infiltration

There was a moderate correlation between the Goutallier classification system and LeanCSA across all levels, both for the left and right sides (left: $R = -0.520$, $P < 0.001$; right: $R = -0.497$, $P < 0.001$; **Figure 5**). There was a strong correlation between Goutallier grades and LeanCSA/CSA across all levels on both sides (left: $R = -0.643$, $P < 0.001$; right: $R = -0.604$, $P < 0.001$; **Figure 5**). Across both sides, aGoutallier correlated strongly with aLeanCSA and aLeanCSA/aCSA ($R = -0.673$ and $R = -0.754$, both $P < 0.001$) (**Supplementary Figure 1**). There was a very strong correlation between the values of the left and right sides for CSA ($R = 0.956$, $P < 0.001$), LeanCSA ($R = 0.900$, $P < 0.001$), and LeanCSA/CSA ($R = 0.827$, $P < 0.001$) considering all levels (**Figure 6**).

Association Between MRI Measurements and Morphometric Parameters

There was a strong correlation between paraspinal muscle CSA and body height (left: $R = 0.737$, $P < 0.001$; right: $R = 0.700$, $P < 0.001$) and between LeanCSA and body height (left: $R = 0.648$, $P < 0.001$; right: $R = 0.612$, $P < 0.001$). There was a weak to moderate correlation between LeanCSA/CSA and body height (left: $R = 0.442$, $P < 0.001$; right: $R = 0.340$, $P = 0.001$). The Goutallier grade correlated weakly with body height ($R = -0.219$, $P = 0.039$). There was a moderate correlation between vertebral body CSA and body height ($R = 0.536$, $P < 0.001$). The paraspinal muscle CSA correlated moderately with the BMI (left: $R = 0.423$, $P < 0.001$; right: $R = 0.436$, $P < 0.001$), and there was no significant correlation between LeanCSA or LeanCSA/CSA and BMI. None of the parameters describing fatty infiltration correlated with age in this cohort.

We observed a moderate correlation between vertebral body CSA and paraspinal muscle CSA (left: $R = 0.448$, $P < 0.001$; right: $R = 0.454$, $P < 0.001$). The vertebral body CSA showed no significant correlation with LeanCSA or LeanCSA/CSA.

DISCUSSION

The aims of this study were to compare paraspinal muscle fatty infiltration as assessed using the Goutallier classification vs. quantitative MRI measurements and to investigate the association between anthropometric parameters and paraspinal muscle morphology and fatty infiltration in patients with sLSS. We observed a moderate to strong correlation between the Goutallier classification system of muscle fatty infiltration and the quantitative assessment of the LeanCSA and lean fraction of the paraspinal muscle CSA, named as LeanCSA/CSA. Moreover, paraspinal muscle morphology and fatty infiltration correlated strongly with body height. These results support the value of assessing fatty infiltration at all levels in the lumbar region and the importance of considering the stature of a patient when interpreting fatty infiltration in the context of LSS.

Patients with higher Goutallier grades had lower LeanCSA and LeanCSA/CSA, representing smaller absolute and relative lean muscle CSA. Similarly, Battaglia et al. reported a strong to very strong correlation for MR images of 25 randomly selected subjects between Goutallier grades and mean percent fat value measured with the same method as in our study

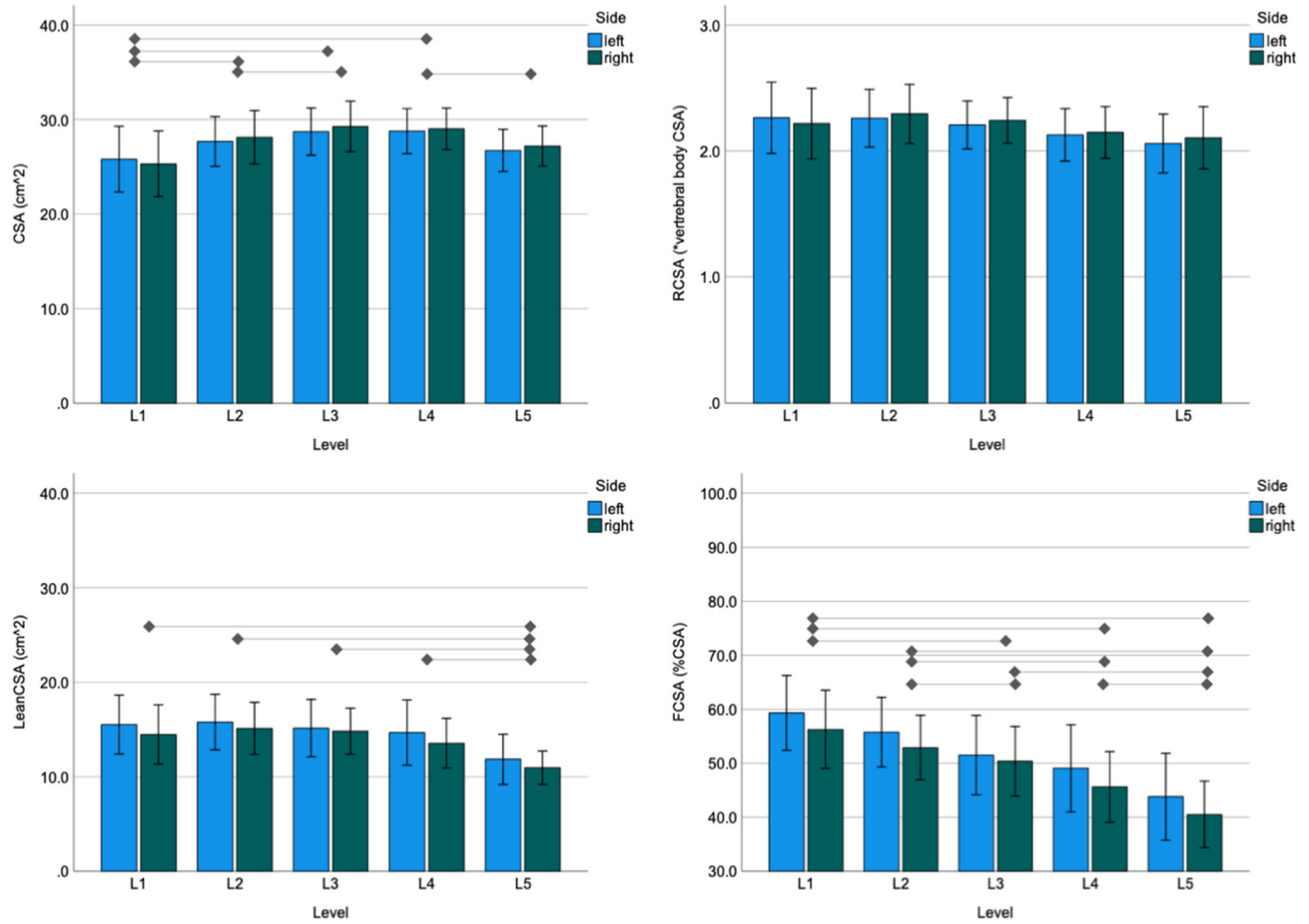


FIGURE 4 | Mean and 95% confidence interval of the paraspinal cross-sectional area (CSA; top left), relative CSA (RCSA; top right), lean muscle cross-sectional area (LeanCSA; bottom left), and the ratio of LeanCSA/CSA and CSA (LeanCSA/CSA; bottom right) for the left and right sides at each level of the lumbar spine. The horizontal lines indicate significant differences between levels. Because significant results of paired comparisons were observed for both sides for each indicated pair, horizontal bars represent differences for both sides (*t*-test for paired samples, $P < 0.05$).

(32). Though, to date, the role of fatty infiltration and atrophy of paraspinal muscles is not clear, there is evidence of their association with LBP—for instance, Kjaer et al. found that fatty infiltration of the multifidus muscle was strongly associated with LBP (5). These results are consistent with those of Goubert et al., who found a positive correlation of fatty infiltration with chronic LBP (4). Other authors investigated the association of muscle morphology and composition with spinal disorders. Yanik et al. reported a significant increase in fatty infiltration in individuals affected by disc herniation (6), and this was corroborated by other studies (15, 41). Yarjanian et al. found a reduced functional CSA of paraspinal muscles in patients with sLSS compared to asymptomatic individuals; however, the difference was not significant when compared to subjects with chronic LBP without evidence of LSS (17). Overall, the evidence for the relation of back muscle to spinal disorders emphasizes the importance for clinicians to regularly evaluate the morphology of paraspinal muscles on the MR images of patients. However, quantitatively measuring paraspinal muscle CSA and lean fraction is time consuming and hence not feasible in the clinical setting. Our results showed that fatty atrophy, assessed using the Goutallier classification system, strongly correlates with quantitative measures and thus—despite being a qualitative measure—is a good and, most importantly, efficient method for evaluating the degeneration of paraspinal muscles in sLSS.

The quantitative measures of paraspinal muscle morphology and fatty infiltration CSA, LeanCSA, and LeanCSA/CSA correlated strongly between the left and right sides across all levels of the lumbar spine. We interpret this as a sign of symmetry in muscle morphology and degeneration. However, because we did not record data on the laterality of symptoms nor of spinal stenosis, we cannot exclude that, in case of asymmetric symptoms or spinal stenosis, an asymmetric muscle degeneration may be observed. In fact, in a study on patients affected by disc herniation, Battié et al. reported greater fatty infiltration in the multifidus muscle on the side of the radicular compression at the level below the herniation, although the total paraspinal muscle CSA was greater on the affected side (15). In another study

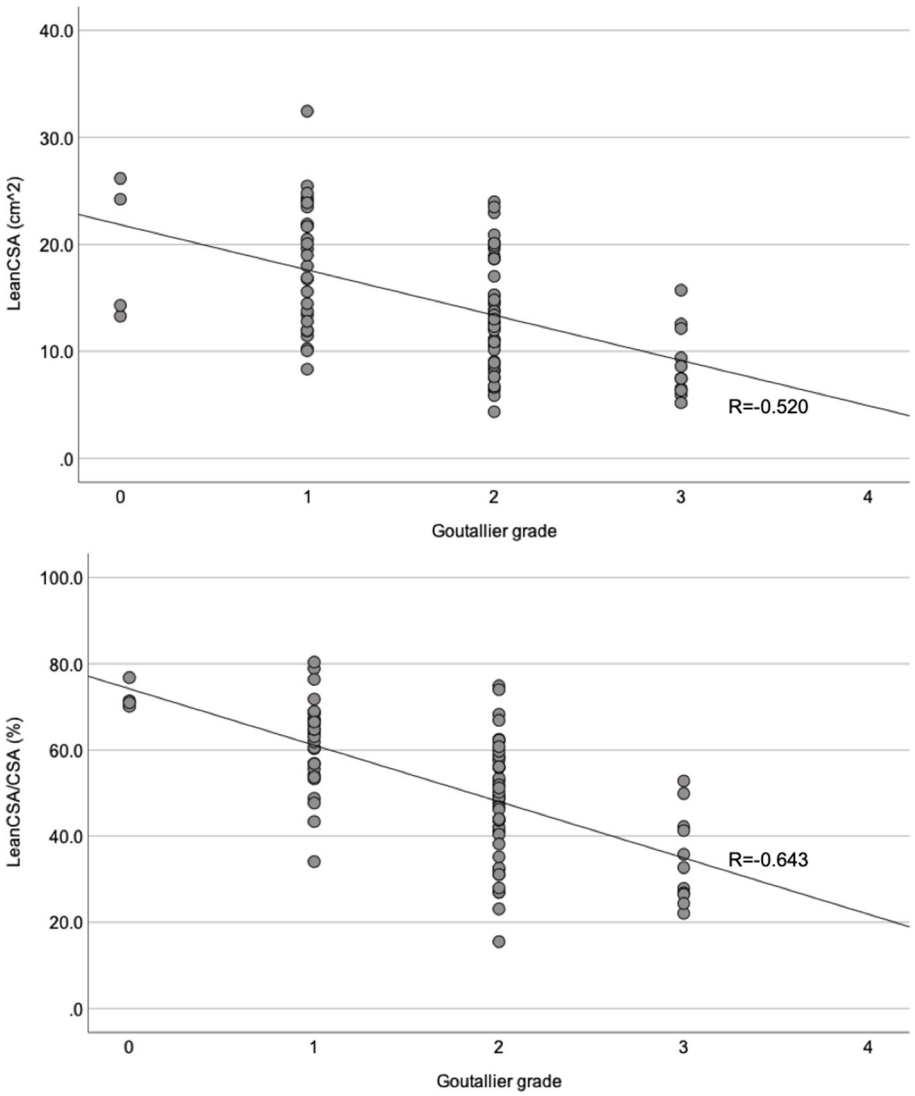


FIGURE 5 | Scatter plots comparing the Goutallier grade for the paraspinal muscles with the lean muscle cross-sectional area (LeanCSA; top) and the ratio between LeanCSA and paraspinal muscle cross-sectional area (LeanCSA/CSA; bottom) of the left side. *R*, Spearman's correlation coefficient.

on patients with symptomatic posterolateral disc herniation at L4 and L5, Fortin et al. did not observe an asymmetry of the multifidus but of the erector spinae muscle, with smaller CSA and greater fatty infiltration on the affected side both of the multifidus muscle and the erector spinae muscle (41). The mechanism of these findings seems to be related to abnormal muscle activation as a result of the altered neural signal (42). Conversely, we did not observe any significant asymmetry of the paraspinal muscles. While we did not assess a possible laterality of symptoms and laterality, we would have only been able to detect a systematic laterality of all patients, which was not the scope of this study. Another possible explanation for this discrepancy is that LSS and its symptoms are more commonly bilateral compared to disc herniation. Nevertheless, further investigation is needed to support this statement.

Fatty infiltration generally increased from cranial to caudal, with the highest value at L5. This result confirms previous studies by Kjaer et al. (5) and Lee et al. (14), who showed that L4 and L5 are the segments most affected by fatty infiltration. In our study, both LeanCSA and LeanCSA/CSA differed significantly between L5 and all other levels. One possible explanation is that most degenerative spinal processes involve the segments L4/L5 and L5/S1 (43), but there is a lack of evidence regarding a causal link between fatty atrophy of paraspinal muscles and spinal disorders. Furthermore, Lee et al. (14) observed an increasing fatty infiltration toward the more caudal levels in 10 healthy volunteers. Kjaer et al. (5) found that the lowest lumbar levels had the greatest signs of fatty infiltration regardless of age in adults and adolescents with a history of LBP. These results suggest that the most caudal lumbar levels are physiologically more

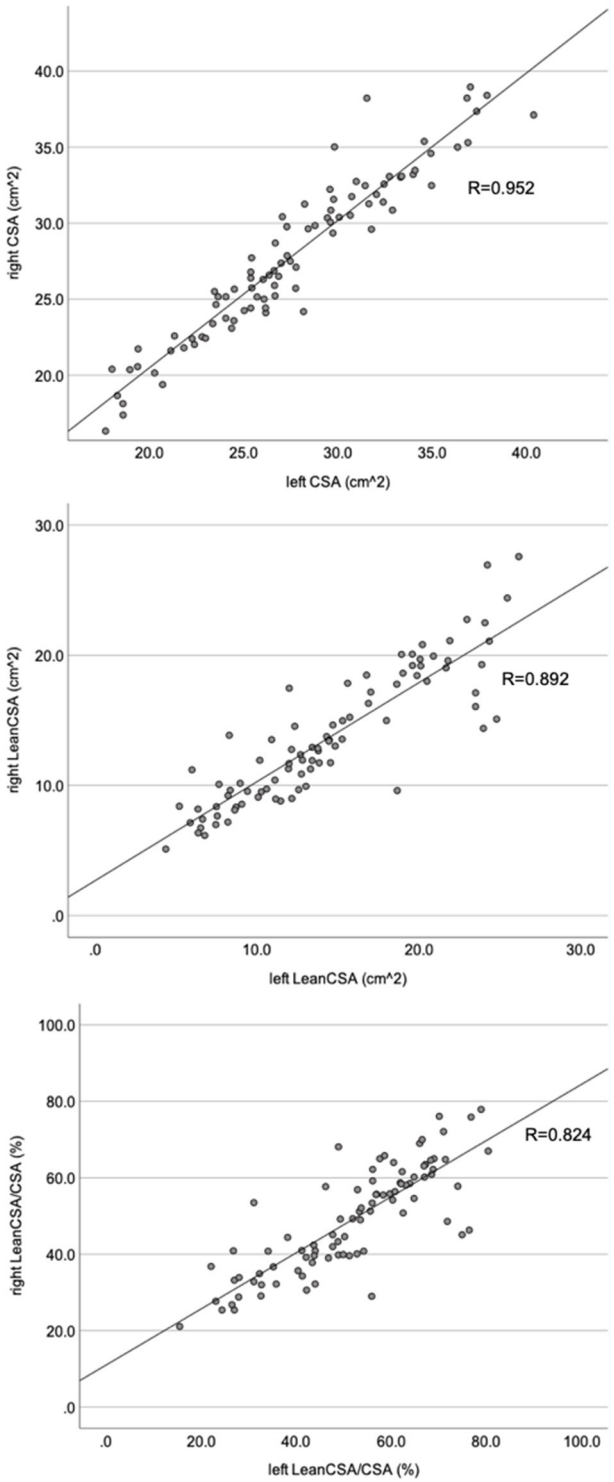


FIGURE 6 | Scatter plots comparing the left and right sides of the paraspinal muscle cross-sectional area (CSA; top), lean muscle cross-sectional area (LeanCSA; middle) and the ratio between LeanCSA and the paraspinal muscle cross-sectional area (LeanCSA/CSA; bottom). *R*, Pearson's correlation coefficient.

prone to fatty infiltration even in young and healthy individuals. Spinal pathologies known to be more frequent at these levels and to be associated with an increase in fatty infiltration of the back muscles may further exaggerate the difference in muscle composition between the cranial and caudal portions of the paraspinal muscles.

Our secondary aim was to determine the association between quantitative measurements of paraspinal muscle morphology and fatty infiltration and anthropometric parameters. We found that CSA and LeanCSA had a strong and LeanCSA/CSA had a weak correlation with body height. The results of previous studies on this association are inconclusive—for instance, our results confirm those reported by Gibbons et al., who observed a positive correlation between the CSA of paraspinal muscles and body height, body mass, and BMI in a large sample of male monozygotic twins (130 subjects) (26). In contrast, two smaller studies by Wood et al. (25) and McGill et al. (23) (13 and 26, respectively) did not find any significant association of CSA with body height, body mass, or BMI in persons with various suspected injuries and diseases. Moreover, Reid et al. reported that greater body height was not a predictor of larger CSA (24). Biomechanically, an association of CSA with body height is intuitive because the greater a person is, the greater are the moment arms of the forces generated by the weight of the upper torso, arms, and head relative to the lumbar spine area during daily activities. Frequently, the paraspinal total and lean CSA are normalized to the vertebral body or intervertebral disc CSA (19, 44–47). It is interesting to note that, in our study, both overall CSA and LeanCSA correlated more strongly with body height than with vertebral CSA (strong *vs*. moderate correlation). Although we consider the vertebral body or the disc as reasonable choice of reference for reducing the bias of differences in anthropometric measures between persons, our results suggest that body height greatly influences paraspinal muscle CSA and should hence be used for normalization.

In our study, only CSA, but not LeanCSA or LeanCSA/CSA, correlated with BMI. This result is consistent with the study by Kjaer et al. (5) who did not find BMI to have an influence on the amount of fatty infiltration in the paraspinal muscles. Moreover, Gibbons et al. (26) found that a greater BMI was a predictor of greater CSA, and Kalichman et al. (48) reported a negative correlation between muscle density on CT scan and BMI. Our results show that people with higher BMI are not prone to greater fatty infiltration of paraspinal muscles, possibly because stronger muscles are required to account for the larger body weight in obese persons when stabilizing the spine. These observations suggest that the higher levels of fatty infiltration of paraspinal muscles are likely not related to obesity but rather to the degenerative process of the muscle itself or to the spinal disorder affecting the patient.

The main strength of our study is the measurement of qualitative and quantitative parameters of paraspinal muscles at each level of the lumbar spine regardless of the level of stenosis. Despite the evidence that fatty infiltration is greater between L4 and S1 and given the anatomy and function of the

erector spinae and multifidus muscles, we believe that these entire muscles should be considered as a single unit. Hence, we introduced global qualitative and quantitative parameters for assessing fatty atrophy of the entire paraspinal muscle unit which aGoutallier and aLeanCSA defined as the average of the respective parameters across all levels. The main limitation of our study is the small sample size, and thus we could not correct for multiple testing. However, while this study can be considered a proof-of-principle investigation, the strong correlations and clear differences observed in our study suggest that these results should hold true when investigated in larger samples. Moreover, we included only individuals affected by sLSS and scheduled for surgery. Similar analyses should be conducted in patients with less severe LSS or who were treated conservatively and in age-matched healthy controls to elucidate the role of muscle atrophy and fatty infiltration in the etiology of the disease and the effects of different treatments on these parameters. The MR images were either obtained at our clinic or transferred by external providers. While using the same make and model for all measurements may have provided even better results, we intended to conduct this study in a real-world environment. In other words, if the same make and model had been used for all patients, the generalizability of the data to other MRI systems would have been unknown/limited. In contrast, our data showed strong correlations despite the different make and models of the MRI systems used, suggesting that the agreement between the assessments is robust regarding the specific MRI system, making our result even more relevant in a clinical context. Finally, while the quantitative method employed here has shown good interobserver and excellent intraobserver reliability, the accuracy of the thresholding procedure has yet to be confirmed with MRI sequences based on chemical shift (Dixon) or spectroscopy. Nonetheless, the sequence used here is a standard clinical sequence, and extracting information on fatty infiltration in standard clinical sequences is relevant especially in cases where Dixon sequences are not available.

CONCLUSION

The correlation of the Goutallier classification with the quantitative assessment of fatty infiltration of paraspinal muscles suggests that clinicians should consider this classification as an efficient tool for evaluating paraspinal muscle fatty infiltration. Nonetheless, this qualitative measure does not consider muscle morphology, which may add insight into the role that paraspinal muscle status plays in the etiology of LSS. Because paraspinal muscle CSA correlates with body height, we suggest taking body height as a reference for normalization in future studies assessing paraspinal muscle atrophy and fatty infiltration.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethikkommission Nordwest- und Zentralschweiz (EKNZ). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

FM: conceptual study design, data collection and processing, and manuscript writing. CNü: conceptual study design, patient recruitment, data collection and processing, statistical analysis, and manuscript review. YZ: data compilation and manuscript review. FH: statistical analysis and manuscript review. SS: conceptual study design and manuscript review. AM: conceptual study design, data analysis, statistical analysis, and manuscript writing and review. CNe: conceptual study design, patient recruitment, and manuscript review. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the Department of Spinal Surgery and the Department of Orthopaedics and Traumatology of the University Hospital Basel, Switzerland.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fneur.2021.656487/full#supplementary-material

Supplementary Figure 1 | Scatter plots comparing the average Goutallier grade across all levels (aGoutallier) with the average lean muscle cross-sectional area (aLeanCSA; top) and the ratio between average LeanCSA and average paraspinal muscle cross-sectional area (aLeanCSA/aCSA; bottom). *R*, Pearson's correlation coefficient.

REFERENCES

1. Lee J-H, Hoshino Y, Nakamura K, Kariya Y, Saita K, Ito K. Trunk muscle weakness as a risk factor for low back pain: a 5-year prospective study. *Spine.* (1999) 24:54–57. doi: 10.1097/00007632-199901010-00013
2. Belavý DL, Armbrecht G, Richardson CA, Felsenberg D, Hides JA. Muscle atrophy and changes in spinal morphology: is the lumbar spine vulnerable after prolonged bed-rest? *Spine.* (2011) 36:137–45. doi: 10.1097/BRS.0b013e3181cc93e8
3. Kalichman L, Carmeli E, Been E. The association between imaging parameters of the paraspinal muscles, spinal degeneration, and low back pain. *BioMed Res Int.* (2017) 2017:1–14. doi: 10.1155/2017/2562957
4. Goubert D, De Pauw R, Meeus M, Willems T, Cagnie B, Schouppe S, et al. Lumbar muscle structure and function in chronic versus recurrent low back pain: a cross-sectional study. *Spine J.* (2017) 17:1285–96. doi: 10.1016/j.spinee.2017.04.025
5. Kjaer P, Bendix T, Sorensen JS, Korsholm L, Leboeuf-Yde C. Are MRI-defined fat infiltrations in the multifidus muscles associated with low back pain? *BMC Med.* (2007) 5:2. doi: 10.1186/1741-7015-5-2

6. Yanik B, Keyik B, Conkbayir I. Fatty degeneration of multifidus muscle in patients with chronic low back pain and in asymptomatic volunteers: quantification with chemical shift magnetic resonance imaging. *Skeletal Radiol.* (2013) 42:771–8. doi: 10.1007/s00256-012-1545-8
7. Kalichman L, Kim DH, Li L, Guermazi A, Hunter DJ. Computed tomography-evaluated features of spinal degeneration: prevalence, intercorrelation, and association with self-reported low back pain. *Spine J.* (2010) 10:200–8. doi: 10.1016/j.spinee.2009.10.018
8. Masaki M, Ikezoe T, Fukumoto Y, Minami S, Tsukagoshi R, Sakuma K, et al. Association of sagittal spinal alignment with thickness and echo intensity of lumbar back muscles in middle-aged and elderly women. *Arch Gerontol Geriatr.* (2015) 61:197–201. doi: 10.1016/j.archger.2015.05.010
9. Jun HS, Kim JH, Ahn JH, Chang IB, Song JH, Kim TH, et al. The effect of lumbar spinal muscle on spinal sagittal alignment. *Neurosurgery.* (2016) 1:847–55. doi: 10.1227/NEU.0000000000001269
10. Barker KL, Shamley DR, Jackson D. Changes in the cross-sectional area of multifidus and psoas in patients with unilateral back pain: the relationship to pain and disability. *Spine.* (2004) 29:E515–9. doi: 10.1097/01.brs.0000144405.11661.eb
11. Ploumis A, Michailidis N, Christodoulou P, Kalaitzoglou I, Gouvas G, Beris A. Ipsilateral atrophy of paraspinal and psoas muscle in unilateral back pain patients with monosegmental degenerative disc disease. *Br J Radiol.* (2011) 84:709–13. doi: 10.1259/bjr/58136533
12. Hyun JK, Lee JY, Lee SJ, Jeon JY. Asymmetric atrophy of multifidus muscle in patients with unilateral lumbosacral radiculopathy. *Spine.* (2007) 32:E598–602. doi: 10.1097/BRS.0b013e318155837b
13. Kim WH, Lee S-H, Lee DY. Changes in the cross-sectional area of multifidus and psoas in unilateral sciatica caused by lumbar disc herniation. *J Korean Neurosurg Soc.* (2011) 50:201. doi: 10.3340/jkns.2011.50.3.201
14. Lee JC, Cha JG, Kim Y, Kim YI, Shin BJ. Quantitative analysis of back muscle degeneration in the patients with the degenerative lumbar flat back using a digital image analysis: comparison with the normal controls. *Spine.* (2008) 33:318–25. doi: 10.1097/BRS.0b013e318162458f
15. Battié MC, Niemelainen R, Gibbons LE, Dhillon S. Is level- and side-specific multifidus asymmetry a marker for lumbar disc pathology? *Spine J.* (2012) 12:932–9. doi: 10.1016/j.spinee.2012.08.020
16. Hicks GE, Simonsick EM, Harris TB, Newman AB, Weiner DK, Nevitt MA, et al. Trunk muscle composition as a predictor of reduced functional capacity in the health, aging and body composition study: the moderating role of back pain. *J Gerontol A Biol Sci Med Sci.* (2005) 60:1420–4. doi: 10.1093/gerona/60.11.1420
17. Yarjanian JA, Fetzer A, Yamakawa KS, Tong HC, Smuck M, Haig A. Correlation of paraspinal atrophy and denervation in back pain and spinal stenosis relative to asymptomatic controls. *PM R.* (2013) 5:39–44. doi: 10.1016/j.pmrj.2012.08.017
18. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. *Spine.* (2005) 30:1441–5; discussion 1446–7. doi: 10.1097/01.brs.0000166503.37969.8a
19. Chen YY, Pao JL, Liaw CK, Hsu WL, Yang RS. Image changes of paraspinal muscles and clinical correlations in patients with unilateral lumbar spinal stenosis. *Eur Spine J.* (2014) 23:999–1006. doi: 10.1007/s00586-013-3148-z
20. Fortin M, Lazáry À, Varga PP, Battié MC. Association between paraspinal muscle morphology, clinical symptoms and functional status in patients with lumbar spinal stenosis. *Eur Spine J.* (2017) 26:2543–51. doi: 10.1007/s00586-017-5228-y
21. Wang W, Sun Z, Li W, Chen Z. The effect of paraspinal muscle on functional status and recovery in patients with lumbar spinal stenosis. *J Orthop Surg Res.* (2020) 15:235. doi: 10.1186/s13018-020-01751-1
22. Zotti MGT, Boas FV, Clifton T, Piche M, Yoon WW, Freeman BJC. Does pre-operative magnetic resonance imaging of the lumbar multifidus muscle predict clinical outcomes following lumbar spinal decompression for symptomatic spinal stenosis? *Eur Spine J.* (2017) 26:2589–97. doi: 10.1007/s00586-017-4986-x
23. McGill SM, Patt N, Norman RW. Measurement of the trunk musculature of active males using CT scan radiography: implications for force and moment generating capacity about the L4/L5 joint. *J Biomech.* (1988) 21:329–41. doi: 10.1016/0021-9290(88)90262-X
24. Reid JG, Costigan PA, Comrie W. Prediction of trunk muscle areas and moment arms by use of anthropometric measures. *Spine.* (1987) 12:273–5. doi: 10.1097/00007632-198704000-00015
25. Wood S, Pearsall DJ, Ross R, Reid JG. Trunk muscle parameters determined from MRI for lean to obese males. *Clin Biomech.* (1996) 11:139–44. doi: 10.1016/0268-0033(95)00018-6
26. Gibbons LE, Videman T, Battié MC, Kaprio J. Determinants of paraspinal muscle cross-sectional area in male monozygotic twins. *Phys Ther.* (1998) 78:602–10; discussion 611–2. doi: 10.1093/ptj/78.6.602
27. Goutallier D, Postel JM, Bernageau J, Lavau L, Voisin MC. Fatty muscle degeneration in cuff ruptures. Pre- and postoperative evaluation by CT scan. *Clin Orthop Relat Res.* (1994):78–83. doi: 10.1097/00003086-199407000-00014
28. Oh JH, Kim SH, Choi J-A, Kim Y, Oh CH. Reliability of the grading system for fatty degeneration of rotator cuff muscles. *Clin Orthop Relat Res.* (2010) 468:1558–64. doi: 10.1007/s11999-009-0818-6
29. Slabaugh MA, Friel NA, Karas V, Romeo AA, Verma NN, Cole BJ. Interobserver and intraobserver reliability of the Goutallier classification using magnetic resonance imaging: proposal of a simplified classification system to increase reliability. *Am J Sports Med.* (2012) 40:1728–34. doi: 10.1177/0363546512452714
30. Lee E, Choi JA, Oh JH, Ahn S, Hong SH, Chai JW, et al. Fatty degeneration of the rotator cuff muscles on pre- and postoperative CT arthrography (CTA): is the Goutallier grading system reliable? *Skeletal Radiol.* (2013) 42:1259–67. doi: 10.1007/s00256-013-1660-1
31. Engelken F, Wassilew GI, Köhlitz T, Brockhaus S, Hamm B, Perka C, et al. Assessment of fatty degeneration of the gluteal muscles in patients with THA using MRI: reliability and accuracy of the Goutallier and quartile classification systems. *J Arthroplasty.* (2014) 29:149–53. doi: 10.1016/j.arth.2013.04.045
32. Battaglia PJ, Maeda Y, Welk A, Hough B, Kettner N. Reliability of the Goutallier classification in quantifying muscle fatty degeneration in the lumbar multifidus using magnetic resonance imaging. *J Manipulative Physiol Ther.* (2014) 37:190–7. doi: 10.1016/j.jmpt.2013.12.010
33. Tamai K, Chen J, Stone M, Arakelyan A, Paholpak P, Nakamura H, et al. The evaluation of lumbar paraspinal muscle quantity and quality using the Goutallier classification and lumbar indentation value. *Eur Spine J.* (2018) 27:1005–12. doi: 10.1007/s00586-018-5485-4
34. Fairbank JC, Pynsent PB. The oswestry disability index. *Spine.* (2000) 25:2940–52; discussion 2952. doi: 10.1097/00007632-200011150-00017
35. Mannion AF, Junge A, Fairbank JC, Dvorak J, Grob D. Development of a German version of the Oswestry Disability Index. Part 1: cross-cultural adaptation, reliability, and validity. *Eur Spine J.* (2006) 15:55–65. doi: 10.1007/s00586-004-0815-0
36. Schizas C, Theumann N, Burn A, Tansey R, Wardlaw D, Smith FW, et al. Qualitative grading of severity of lumbar spinal stenosis based on the morphology of the dural sac on magnetic resonance images. *Spine.* (2010) 35:1919–24. doi: 10.1097/BRS.0b013e3181d359bd
37. Fortin M, Battie MC. Quantitative paraspinal muscle measurements: inter-software reliability and agreement using OsiriX and ImageJ. *Phys Ther.* (2012) 92:853–64. doi: 10.2522/ptj.20110380
38. Ranson CA, Burnett AF, Kerslake R, Batt ME, O'Sullivan PB. An investigation into the use of MR imaging to determine the functional cross sectional area of lumbar paraspinal muscles. *Eur Spine J.* (2006) 15:764–73. doi: 10.1007/s00586-005-0909-3
39. Fortin M, Omidyeganeh M, Battie MC, Ahmad O, Rivaz H. Evaluation of an automated thresholding algorithm for the quantification of paraspinal muscle composition from MRI images. *Biomed Eng Online.* (2017) 16:61. doi: 10.1186/s12938-017-0350-y
40. Campbell MJ. *Statistics at Square Two*. 2nd edn. London: Blackwell. BMJ Books. (2006).
41. Fortin M, Lazáry À, Varga PP, McCall I, Battié MC. Paraspinal muscle asymmetry and fat infiltration in patients with symptomatic disc herniation. *Eur Spine J.* (2016) 25:1452–9. doi: 10.1007/s00586-016-4503-7
42. Hodges P, Holm AK, Hansson T, Holm S. Rapid atrophy of the lumbar multifidus follows experimental disc or nerve root injury. *Spine.* (2006) 31:2926–33. doi: 10.1097/01.brs.0000248453.51165.0b
43. Saleem S, Aslam HM, Rehmani MA, Raees A, Alvi AA, Ashraf J. Lumbar disc degenerative disease: disc degeneration symptoms and magnetic resonance image findings. *Asian Spine J.* (2013) 7:322–34. doi: 10.4184/asj.2013.7.4.322

44. Leng J, Han G, Zeng Y, Chen Z, Li W. The effect of paraspinal muscle degeneration on distal pedicle screw loosening following corrective surgery for degenerative lumbar scoliosis. *Spine.* (2020) 45:590–8. doi: 10.1097/BRS.0000000000003336
45. Kim JY, Ryu DS, Paik HK, Ahn SS, Kang MS, Kim KH, et al. Paraspinal muscle, facet joint, and disc problems: risk factors for adjacent segment degeneration after lumbar fusion. *Spine J.* (2016) 16:867–75. doi: 10.1016/j.spinee.2016.03.010
46. Xia W, Fu H, Zhu Z, Liu C, Wang K, Xu S, et al. Association between back muscle degeneration and spinal-pelvic parameters in patients with degenerative spinal kyphosis. *BMC Musculoskelet Disord.* (2019) 20:454. doi: 10.1186/s12891-019-2837-0
47. Urrutia J, Besa P, Lobos D, Campos M, Arrieta C, Andia M, et al. Lumbar paraspinal muscle fat infiltration is independently associated with sex, age, and inter-vertebral disc degeneration in symptomatic patients. *Skeletal Radiol.* (2018) 47:955–61. doi: 10.1007/s00256-018-2880-1
48. Kalichman L, Hodges P, Li L, Guermazi A, Hunter DJ. Changes in paraspinal muscles and their association with low back pain and spinal degeneration: CT study. *Eur Spine J.* (2010) 19:1136–44. doi: 10.1007/s00586-009-1257-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership