Eur Radiol (2014) 24:1610–1620
DOI 10.1007/s00330-014-3145-6

MAGNETIC RESONANCE

Reproducibility, and age, body-weight and gender dependency of candidate skeletal muscle MRI outcome measures in healthy volunteers

Jasper M. Morrow · Christopher D. J. Sinclair · Arne Fischmann · Mary M. Reilly · Michael G. Hanna · Tarek A. Yousry · John S. Thornton

Received: 7 November 2013 / Revised: 12 February 2014 / Accepted: 5 March 2014 / Published online: 20 April 2014

Abstract

Objectives Quantitative magnetic resonance imaging (MRI) can potentially meet the pressing need for objective, sensitive, reproducible outcome measures in neuromuscular disease trials. We tested, in healthy volunteers, the consistency, reliability and sensitivity to normal inter-subject variation of MRI methods targeted to lower limb muscle pathology to inform the design of practical but comprehensive MRI outcome measure protocols for use in imminent patient studies.

Methods Forty-seven healthy volunteers, age 21-81 years, were subject at 3T to three-point Dixon fat-fraction measurement, T_1-relaxometry, T_2-relaxometry and magnetisation transfer ratio (MTR) imaging at mid-thigh and mid-calf level bilaterally. Fifteen subjects underwent repeat imaging at 2 weeks.

Results Mean between-muscle fat fraction and T_2 differences were small, but significant ($p<0.001$). Fat fraction and T_2 correlated positively, and MTR negatively with subject age in both the thigh and calf, with similar significant correlations with weight at thigh level only ($p<0.001$ to $p<0.05$). Scan-rescan and inter-observer intra-class correlation coefficients ranged between 0.62-0.84 and 0.79-0.99 respectively.

Conclusions Quantitative lower-limb muscle MRI using readily implementable methods was sensitive enough to demonstrate inter-muscle differences (small in health), and correlations with subject age and weight. In combination with high reliability, this strongly supports the suitability of these methods to provide longitudinal outcome measures in neuromuscular disease treatment trials.

Key points

- *Quantitative lower limb muscle MRI provides potential outcome measures in neuromuscular diseases*
- *Bilateral thigh/calf coverage using sequences sensitive to acute and chronic pathology*
- *Measurements have excellent scan-rescan and interobserver reliability*
- *Measurements show small but significant inter-subject age and weight dependency*
- *Readily implementable sequences suitable for further assessment in patient studies*

Keywords Magnetic resonance imaging · Neuromuscular diseases · Outcome assessment · Lower extremity · Reference values

Abbreviations

NMD	Neuromuscular disease
MTR	Magnetisation transfer ratio
FF	Fat fraction
ROI	Region of interest
ICC	Intra-class correlation coefficient
H1-MRS	Proton magnetic resonance spectroscopy
TSE	Turbo spin echo

Electronic supplementary material The online version of this article (doi:10.1007/s00330-014-3145-6) contains supplementary material, which is available to authorized users.

J. M. Morrow (✉) · C. D. J. Sinclair · M. M. Reilly · M. G. Hanna · T. A. Yousry · J. S. Thornton
Medical Research Council Centre for Neuromuscular Diseases, Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK
e-mail: j.morrow@ucl.ac.uk

A. Fischmann
Department of Radiology, Division of Diagnostic and Interventional Neuroradiology, University of Basel Hospital, Basel, Switzerland

C. D. J. Sinclair · T. A. Yousry · J. S. Thornton
Neuroradiological Academic Unit, Department of Brain Repair and Rehabilitation, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK

Introduction

Imminent clinical treatment trials for neuromuscular diseases (NMDs) [1, 2] need valid, sensitive and reliable treatment response measures [3]. Conventional outcome measures, including muscle strength, neurophysiology and functional assessment are insufficiently sensitive [2]: NMDs typically progress slowly against a background of age-related changes [4], with therapies more likely to reduce progression than reverse established injury. There is a pressing need for outcome measures reflecting underlying pathological processes with demonstrable longitudinal sensitivity and applicability in multi-centre trials. Systematic assessment of feasibility, reproducibility and normal variation in healthy volunteers is a logical first step in establishing such measures.

Conventional magnetic resonance imaging (MRI) can delineate both acute and chronic muscle pathology: acute denervation [5] and inflammation [6] cause oedema-related T_2-weighted hyper-intensity, typically in early disease and potentially reversible with treatment [7]. Chronic muscle damage, whether caused by a primary myopathy or secondary to a neuropathy, results in atrophy and fatty degeneration [8, 9], causing T_1-weighted hyper-intensity, with patterns aiding NMD diagnosis [9].

Quantitative MRI can objectively measure these changes: on T_2 relaxometry muscle T_2 is elevated in myotonic dystrophy [10], Duchenne muscular dystrophy [11], juvenile dermatomyositis [6] and amyotrophic lateral sclerosis [12], while skeletal muscle magnetisation transfer ratio (MTR) is decreased in limb girdle muscular dystrophy [13] and Charcot-Marie-Tooth disease [14]. Muscle fat content has been quantified by T_1-relaxometry [15], proton spectroscopy (^{1}H-MRS) [16–19], T_2 relaxation modelling [20–22] and chemical-shift based Dixon fat-water separation [23] providing maps of the proportion of fat to water, or "fat fraction" (FF) [10, 22, 24–26].

While these reports support the validity of putative NMD MRI outcome measures, little has been published on potentially confounding age, gender or body mass dependencies [17, 25, 27], and while inter- or intra-observer reproducibility has been investigated [10, 15, 28], scan-rescan reproducibility has not been addressed systematically. These factors significantly influence trial statistical sensitivity [3]. Furthermore, studies seldom compare multiple MRI measures in the same subjects, and at both calf- and thigh-levels, having focused mainly on single measures [6, 13, 14, 21, 27, 29, 30], in either lower leg [10, 13, 21, 26] or thigh [6, 31], and generally in a single limb rather than bilaterally.

To establish practical NMD MRI trial outcome-measure protocols, we assessed in healthy volunteers a suite of MRI measures expected to be sensitive to NMD muscle pathology. We tested: *reproducibility* by quantifying scan-rescan and inter-observer reliability, *internal consistency* by comparing left- and right-limb values, *external consistency* by comparison with published data and *sensitivity* to healthy variation by measuring the dependence of lower-limb muscle T_1, pseudo-T_2, FF and MTR upon anatomical location and demographic factors, including sex, age and body mass.

Materials and methods

Subjects and MRI examination schedule

With local research ethics committee approval and written consent, 47 healthy volunteers (23 men) were studied: (mean ± SD, range) age 44.4±17.0, 21.5-81.0 years; height 171±9, 150-188 cm; weight 73±16, 44-115 kg; body mass index 25±4.7, 17-41 kg/m^2; 15 undergoing repeat imaging after approximately 2 weeks with identical imaging parameters. The subjects, recruited from friends and family of patients participating in MRI research or the host institution staff, underwent clinical screening to exclude neuromuscular disease prior to examination.

MRI sequence selection

Four MRI measures were chosen for investigation according to their likely sensitivity to both acute (T_2 and MTR) and chronic (T_1 and fat fraction) muscle pathology. Specific measurement pulse sequences and parameters were selected on a pragmatic basis: we chose to select from standard pulse sequences widely available on routinely available imaging platforms with imaging parameters selected to facilitate accurate quantification. For the purposes of the present study it was necessary to obtain wide anatomical coverage of both limbs at thigh and calf level in a practical examination time: this necessitated certain compromises in the acquisition design, such as precluding the use of a Carr-Purcell-Meiboom-Gill multi-echo T_2-measurement sequence.

MRI acquisition

Subjects were examined lying feet-first and supine at 3T (TIM Trio; Siemens, Erlangen, Germany) using a multi-channel peripheral angiography coil (PA Matrix; Siemens) and 'spine matrix' coil elements. Before examination, the distance between the anterior superior iliac spine and the superior border of the patella was measured and thigh-level imaging volumes were centred one-third of this distance above the patella superior border. Calf-level imaging volumes were centred on the point of widest lower leg circumference.

Axial-slice matrices and fields of view (FOVs) were 256×128 and 400×200 mm (410×205 mm in some subjects) for thigh-level images and 256×120 and 400×188 mm for calf-level images, except for FF acquisitions where matrices were

512×256 and 512×240 pixels respectively. In this healthy volunteer study, fat suppression was not applied in any of the measurements. The total acquisition time was less than 40 min and included the following sequences:

Fat fraction measurement

For Dixon FF measurement [23], three 2D gradient-echo acquisitions were performed with echo-times (TEs) (TE_1/TE_2/TE_3=3.45/4.60/5.75 ms, TR=100 ms, flip angle [α]=10°, bandwidth [BW]=420 Hz/pixel, number of excitations [NEX]=4, 10×10-mm slices with 10-mm gap). The maps of the field error term, φ, generated as an intermediate step in Glover and Schneider's decomposition algorithm [23], underwent phase unwrapping using the PRELUDE tool, which is part of the FSL software (FMRIB, Oxford) [32]. Each limb in the FOV was processed separately on a 2D individual-slice basis using the TE=3.45 ms magnitude image as a threshold mask. The decomposed fat (*F*) and water (*W*) images were then used to calculate FF as $FF = 100\,\% \times F/(F+W)$. The TE=3.45 ms image was used for region of interest (ROI) placement and as a reference for inter-method image registration using FLIRT (FSL, FMRIB, Oxford).

T_1-relaxometry

DESPOT-1 [33] T_1-mapping used three 3D fast low-angle shot (3D-FLASH) images $S_{1,2,3}$ with nominal $\alpha_{1,2,3}$ of 5, 15 and 25°, TR/TE=23/3 ms, and BW=440 Hz/pixel acquired in a single, non-selective slab with 80×5 mm longitudinal phase-encoded partitions. Flip-angles were corrected using B_1 maps obtained as below and T_1 calculated according to Deoni et al. [33].

T_2-relaxometry

Dual-contrast turbo-spin-echo (TSE) images (TR/TE_1/TE_2= 5,500/16/64 ms, 6,500/13/52 ms or 6,500/16/56 ms; 10×10-mm slices with 10-mm gap, parallel imaging factor (iPat) 2, TSE factor 4, BW=444 Hx/pixel, refocusing flip angle 180º, NEX=2) were acquired. Pseudo-T_2 was calculated from the respective pixel intensities I_{TE1} and I_{TE2} from the TE_1 and TE_2 images as $T_2 = \frac{TE_2 - TE_1}{\ln(I_{TE1}\ I_{TE2})}$.

B_1 mapping

Separate TSE images (TR/TE=7,000/11 ms, 128×64 matrix, 40 contiguous 10-mm slices, BW=429 Hx/pixel, 1/2*k*-space sampling) yielded image intensities V_1 and V_2 acquired with nominal excitation α_1 and α_2 of 60° and 120°. B_1 deviation was mapped according to $B_{1Dev}=\arccos(V_2/2V_1)/\alpha_1$ [34].

Magnetisation transfer ratio

MTRs were calculated from two 3D-FLASH images with (M_1) and without (M_0) an MT pre-pulse (500° amplitude, 1,200 Hz offset, 10 ms duration) (TR/TE=65/3 ms or 68/3 ms, α=10º, BW=440 Hz/pixel, iPat=2, 40×5-mm longitudinal phase encoding partitions) according to $MTR=(M_0-M_1)/M_0\times 100$ percentage units (p.u.). MTR maps were RF-inhomogeneity corrected using B_1 maps obtained as described in "B1 mapping" above according to [35] using a mean-over-all-subjects B_1 inhomogeneity correction factor of k=0.0085.

ROI analysis

A single observer (A.F.; a radiologist with 4 years post-specialist experience in neuromuscular imaging) defined ROIs outlining the cross-sectional area of each muscle avoiding contamination with fascia or subcutaneous and inter-muscular fat and allowing for minor movement between acquisitions, using ITK-SNAP [36]. The fifth-most superior slice was used in the thigh and the sixth slice in the calf, unless muscles below were not visible, in which case an adjacent slice was selected. In the 15 subjects with repeated imaging, ROIs for the second acquisition were drawn on the slice most similar to that used from the first acquisition.

Left and right limb ROIs were defined for the rectus femoris, vastus lateralis, vastus intermedius, vastus medialis, semimembranosus, semitendinosus, biceps femoris, adductor magnus, sartorius, gracilis, tibialis anterior, peroneus longus, lateral gastrocnemius, medial gastrocnemius, soleus and tibialis posterior muscles (Fig. 1a). The ROIs were transferred to the co-registered parameter maps, minor position adjustments to account for imperfect registration were performed as necessary and the mean value for each muscle ROI was recorded. To provide summary measures, the mean of all individual-muscle ROI-means for each subject was calculated for each measure separately at thigh and at calf level. To assess inter-observer reliability, a second observer (J.M.; a neurologist with 3 years' experience in neuromuscular imaging) independently defined ROIs using the same method on one acquisition from each of the 15 subjects with repeat examinations. Image data were inspected visually and ROI values originating from areas of gross artefact were excluded from the analysis.

Statistics

Using SPSS 18 (SPSS, Chicago, IL), inter-muscle differences were assessed using ANOVA with post hoc comparisons using Bonferroni's method. Inter-scan and inter-observer overall mean value differences were assessed using paired *t*-tests and reproducibility determined as mean absolute inter-scan and inter-observer differences, displayed on Bland-Altman plots with calculation of limits of agreement [37]

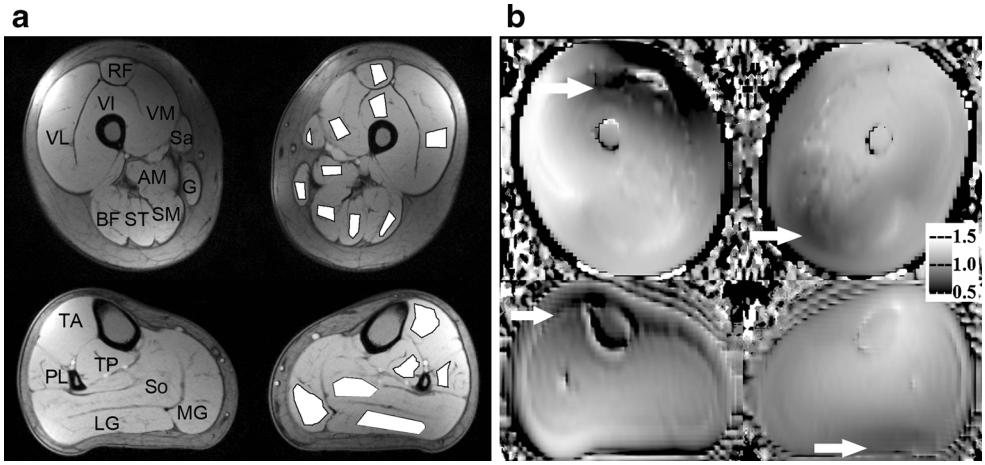

Fig. 1 Sample images from a single volunteer (a 24-year-old man, both thighs and calves). **a** Unprocessed Dixon acquisition (TE=3.45 ms) used for definition of ROIs demonstrated on left thigh and calf. **b** B_1 field map demonstrating reduced B_1 anteriorly on right and posteriorly on left (*arrows*). All images are axial with standard orientation (anterior at top of image, subject's right hand side at left of image). ROI labels in the thigh: *RF* rectus femoris, *VM* vastus medialis, *VI* vastus intermedius, *VL* vastus lateralis, *Sa* sartorius, *SM* semimembranosus, *ST* semitendinosus, *BF* biceps femoris (long head), *AM* adductor magnus, *G* gracilis. ROI labels in the calf: *TA* tibialis anterior, *TP* tibialis posterior, *PL* peroneus longus, *So* soleus, *MG* medial head of gastrocnemius, *LG* lateral head of gastrocnemius

and intra-class correlation coefficients (ICCs). Multivariate regression assessed the influence of demographic factors (age, gender, weight, height) on MRI measures: height showed no independent correlation with any MRI measure and was therefore excluded from the model. Pearson's correlation coefficients between MRI measures were calculated.

Results

Data quality

The number of images excluded from the analysis was small: nine data-sets were missing or technically non-analysable: FF—thigh 1, calf 1; T_1—thigh 2, calf 4; T_2—none, MTR—calf 1. In the remaining data, small fractions of individual ROIs were excluded due to local artefact, mostly B_1-related signal drop-out: FF—thigh 1.7 % (16/920), calf 2.4 % (13/540); T_1—thigh 24 % (219/900), calf 12 % (57/492); T_2—thigh 5.4 % (51/940), calf 0.2 % (1/552), MTR—thigh 15 % (142/920), calf 5.2 % (28/540). In all subjects, asymmetric B_1 deviations were observed (Fig. 1b) with B_1 reduced anteriorly on the right and posteriorly on the left. This was evident at the calf level but more prominent in the thigh, particularly affecting the right rectus femoris and vastus medialis. This artefact prevented measurement within right rectus femoris in 45/47 T1 maps, 41/47 MTR maps and within right vastus medialis in 35/47 T1 maps and 33/47 MTR maps.

Individual muscle values

MRI parameter maps from a representative subject are depicted in Fig. 2. Individual muscle values for each MRI measure in all 47 subjects are shown in Fig. 3. FF and T_2 were similar in the left and right limbs, suggesting asymmetric B_1 variations did not unduly influence these measures. Between muscles, FF differed significantly (ANOVA, $p<0.001$ at both calf- and thigh-level). Group-mean sartorius FF was higher than all other thigh-level muscles ($p<0.01$ for semimembranosus, $p<0.001$ for all other muscles), whilst the rectus femoris FF was lower than most other thigh muscles ($p<0.01$ vs gracilis, vastus lateralis; $p<0.001$ vs sartorius, semimembranosus, biceps femoris and adductor magnus). Similarly, in the calf soleus the FF was highest ($p<0.05$ vs peroneal, $p<0.01$ vs medial gastrocnemius, $p<0.001$ vs each remaining muscle), whilst tibialis anterior FF was the smallest ($p<0.01$ vs medial and lateral gastrocnemius, $p<0.001$ vs soleus and peroneal). However, the absolute inter-muscle differences were small; FF ranging from 0.6 % in the rectus femoris to 2.9 % in the sartorius. Inter-muscle T_2 differences were also significant (ANOVA, $p<0.001$ at both calf and thigh-level), with the same muscles (sartorius, semimembranosus and biceps femoris in the thigh; soleus, peroneal in the calf) showing elevated T_2 as elevated FF. Whilst tibialis posterior and tibialis anterior T_2 times were lowest in the calf, consistent with their low FF, gracilis T_2 was lowest despite this muscle's intermediate FF.

MTR showed apparent left-right differences in some regions with lower values for right tibialis anterior, right rectus femoris and left semimembranosus, corresponding to the areas of maximum B_1 deviation. Excepting these ROIs, MTR was similar across all thigh and calf muscles (range, 31.7-33.2 p.u.). Mean T_1 similarly varied between left and right limbs in these muscles suggesting incomplete B_1 inhomogeneity correction, but was otherwise consistent across the remaining muscles (1,240-1,370 ms).

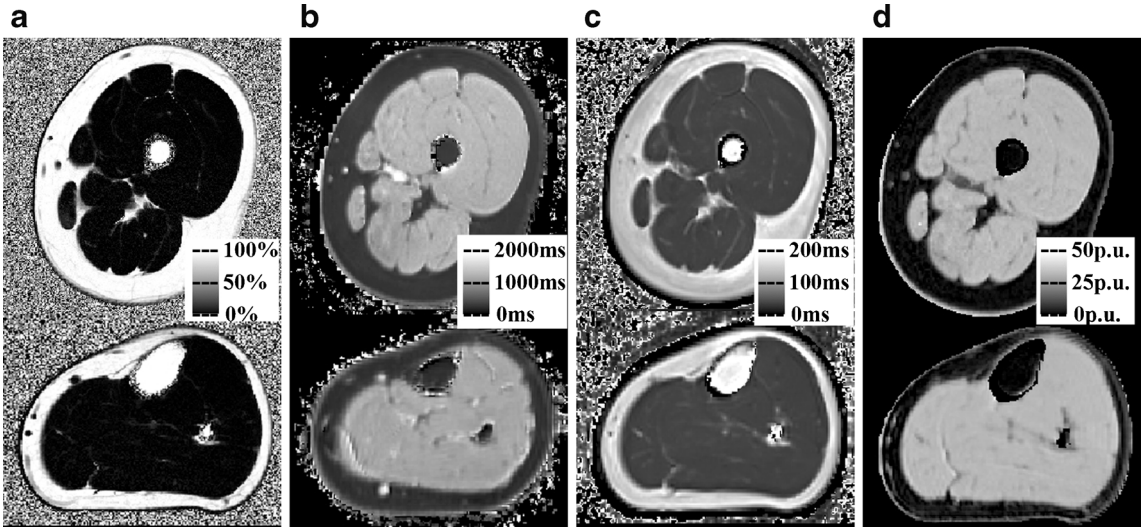

Fig. 2 Sample quantitative maps from a single volunteer (a 24-year-old man, left thigh and calf). **a** Fat fraction map (in %). **b** T_1 map in ms at left thigh and calf level. **c** T_2 map (in ms). **d** MTR map (in p.u.). All images are axial with standard orientation (anterior at top of image, right hand side at left of image)

Scan-rescan and inter-observer reliability

Scan-rescan reliability values are shown in Table 1, with inter-observer reliability in Table 2. Mean values are shown for both summary measures and individual-muscle ROI values, together with scan-rescan and inter-observer ICCs and limits of agreement for both. ICCs were 0.84-0.99 for inter-observer and 0.62-0.99 for scan-rescan values, and were generally higher for the summary measures than for the individual muscle values. The limits of agreement were consistently narrower for overall mean values and inter-observer comparisons than for individual ROI values and inter-scan comparisons. The limits of agreement were broadly similar when each muscle was analysed separately (ESM Table 1).

Dependence upon age, gender and weight

Results of multivariate linear regression modelling the MRI measures at each level against the assumed explanatory variables age, gender and weight are shown in Table 3 for the all-muscle summary measures, and for individual muscles in ESM Table 2. There were significant positive dependencies of both FF and T2 upon age at both anatomical levels, and upon weight in the thigh but not calf. MTR showed strong negative dependence upon age ($p<0.001$) for both thigh and calf (see also Fig. 4, illustrating the univariate Pearson correlation between overall muscle mean MTR and age), and significant correlation with weight and notably gender in the thigh. T_1 did not depend significantly upon any demographic parameter, except for an association with weight in the thigh only ($p<0.05$). Although FF correlated positively with T_2, and negatively with T_1 and MTR (Table 4), the MTR-age correlation remained significant when the other quantitative parameters were included as covariates ($p<0.01$ thigh, $p<0.001$ calf). We also constructed multivariate linear regression models for individual muscles (ESM Table 2), most consistently demonstrating positive correlations between FF or T_2 and weight in the thigh, and negative correlations between MTR and age/gender/weight in the thigh, and age in the calf.

Discussion

We demonstrated the reproducibility of 3T MRI lower limb muscle T_1, T_2, MTR and FF obtained using routinely available acquisition sequences suitable for deployment in NMD treatment trials. With the exception of T_1 and MTR in areas of poor B_1 homogeneity, we obtained literature-consistent measurements with good internal consistency, and demonstrated dependence upon specific muscle compartment, age and weight in healthy individuals. Since changes in these measures with muscle disease are expected to far exceed the variations in health we report, combinations of these measurements targeted to disease-specific anatomical levels may offer robust trial outcome measures sensitive to pathological change.

Inter-muscle variation and comparison with previous studies

We observed small but significant inter-muscle T_2 and FF differences, including hamstring FF exceeding quadriceps FF [22], and increased soleus T_2 compared with tibialis anterior, consistent with previous results [22, 25, 26, 38] attributed

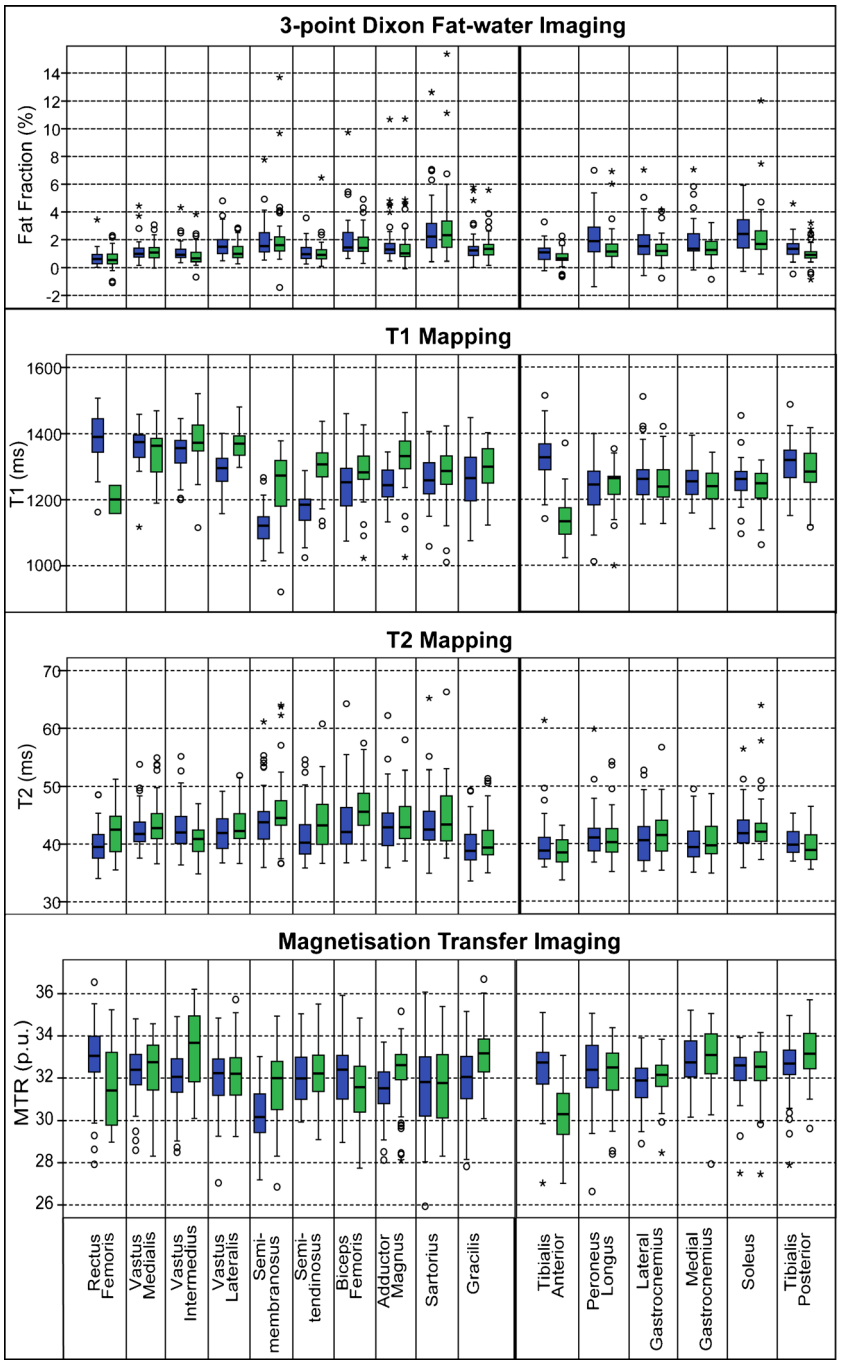

Fig. 3 Individual muscle ROI values at thigh and calf levels for 47 subjects. *Bars* indicate median, 25th, 50th and 75th centiles, *blue* left limb, *green* right limb, *lines* range, ° minor outlier, * major outlier. *MTR* magnetisation transfer ratio, *p.u.* percentage units

to differing proportions of type 1 muscle fibres [39] with increased intra-myocellular lipid [38]. For outcome assessment, this anatomical specificity far exceeds that provided by non-imaging outcome measures such as myometry [40] and neurophysiology [41]. Excepting those muscles for which B_1 deviations were too severe for effective correction, MTRs were consistent with previous calf-muscle studies [13, 25]. All measurements showed good left-right internal consistency except T_1 and MTR in areas of maximum B_1 variation where correction was impossible or proved inadequate.

Reproducibility

The inter-scan limits of agreement provide a measure of sensitivity to detect meaningful change; e.g. for the thigh-level, a change in the overall mean measures in FF, T_2, T_1 or MTR of +0.28 %, +1.8 ms, -39 ms or -1.63 p.u. is a significant change at the 95 % level for an individual subject. Rates of change of these with specific NMD progression will be confirmed in future natural history studies, but the detectable change thresholds our data suggest are small compared with

Table 1 Inter-scan reliability of MRI measurements from ROIs defined by a single observer for both summary measures and individual muscle ROI values

Measure	1st scan group (mean ± SD)	2nd scan group (mean ± SD)	ICC	Limits of agreement	*n*
Thigh level—mean across all ROIs for each subject					
Fat-Fraction (%)	1.36±0.50	1.25±0.58	0.91	-0.51 to +0.28	14
T_1 (ms)	1,290±32	1,288±30	0.65	-39 to +35	14
T_2 (ms)	42.01±2.28	42.19±2.29	0.94	-1.47 to +1.83	14
MTR (p.u.)	32.23±1.40	32.25±1.19	0.87	-1.63 to +1.67	14
Calf level—mean across all ROIs for each subject					
Fat-Fraction (%)	1.54±0.65	1.30±0.56[a]	0.89	-0.58 to +0.08	15
T_1 (ms)	1,276±66	1,283±56	0.62	-100 to +114	13
T_2 (ms)	39.89±1.75	40.23±2.14	0.83	-1.84 to +2.54	15
MTR (p.u.)	32.80±0.57	32.91±0.44	0.69	-0.67 to +0.89	14
Thigh level—individual ROI values					
Fat-Fraction (%)	1.32±0.87	1.20±0.82[a]	0.76	-1.25 to +1.01	254
T_1 (ms)	1,282±88	1,286±81	0.79	-103 to +111	190
T_2 (ms)	41.90±3.17	42.11±3.27	0.83	-3.43 to +3.83	255
MTR (p.u.)	32.35±1.79	32.40±1.47	0.71	-2.39 to +2.49	224
Calf level individual ROI values					
Fat-Fraction (%)	1.51±0.93	1.21±0.85[a]	0.62	-1.55 to +0.95	166
T_1 (ms)	1,271±83	1,281±79	0.65	-122 to +142	119
T_2 (ms)	39.89±2.60	40.23±2.92[a]	0.79	-3.11 to +3.80	180
MTR (p.u.)	32.82±1.35	32.89±1.03	0.65	-1.9 to +2.04	160

Limits of agreement are calculated±1.96 SD by Bland-Altman method

ICC intra-class correlation coefficient

[a] Evidence of systematic difference between scans ($p<0.001$, Bland-Altman method)

cross-sectional disease-dependencies [10, 14, 15, 21, 22] and are in the range of 1-year changes in oculopharyngeal muscle dystrophy [42].

Inter-scan differences exceeded inter-observer differences as a source of variation, the former potentially driven by small scan-scan position inconsistencies. Compliance with a

Table 2 Inter-observer reliability of MRI measurements from identical source images for both summary measures and individual muscle ROI values

Measure	1st observer group (mean ± SD)	2nd observer group (mean ± SD)	ICC	Limits of agreement	*n*
Thigh level—overall mean for each subject					
Fat fraction (%)	1.33±0.50	1.24±0.41[a]	0.93	-0.39 to +0.21	15
T_1 (ms)	1,293±33	1,287±34	0.95	-26 to +14	15
T_2 (ms)	41.97±2.20	42.01±2.22	0.98	-0.93 to +1.00	15
MTR (p.u.)	32.29±1.37	32.34±1.27	0.99	-0.49 to +0.59	15
Calf level—overall mean for each subject					
Fat fraction (%)	1.54±0.65	1.51±0.57	0.95	-0.41 to +0.33	15
T_1 (ms)	1,275±63	1,276±66	0.99	-23 to +24	14
T_2 (ms)	39.89±1.75	39.75±1.79	0.96	-1.08 to +0.80	15
MTR (p.u.)	32.75±0.57	32.84±0.56	0.95	-0.25 to +0.43	15
Thigh level—individual ROI values					
Fat fraction (%)	1.32±0.91	1.23±0.91[a]	0.79	-1.04 to +1.22	281
T_1 (ms)	1,289±88	1,284±94	0.93	-62 to +72	221
T_2 (ms)	41.96±3.19	42.00±3.61	0.84	-3.82 to +3.75	280
MTR (p.u.)	32.36±1.80	32.34±1.91	0.90	-1.59 to +1.63	248
Calf level individual ROI values					
Fat fraction (%)	1.55±0.97	1.50±0.98	0.83	-1.08 to +1.17	172
T_1 (ms)	1,269±81	1,269±81	0.95	-48 to +50	138
T_2 (ms)	39.89±2.60	39.75±2.69	0.86	-2.59 to +2.88	180
MTR (p.u.)	32.76±1.36	32.83±1.31	0.92	-1.15 to +1.00	175

Limits of agreement are calculated±1.96 SD by Bland-Altman method

ICC intra-class correlation coefficient

[a] Evidence of systematic difference between scans/observers ($p<0.001$, Bland-Altman method)

Table 3 Multivariate regression analysis of the dependence of mean muscle MRI measures in thigh and calf upon demographic factors in healthy volunteers

	Thigh		Calf	
FF	R=0.58, p=0.001		R=0.42, p<0.05	
	Co-eff	p	Co-eff	p
Constant	-0.942	0.150	0.573	0.399
Gender	-0.125	0.597	-0.329	0.174
Age	*0.016*	*0.026*	*0.014*	*0.047*
Weight	*0.025*	*0.003*	0.008	0.374
T_1	R=0.50, p=0.01		R=0.43, p=0.10	
	Co-eff	p	Co-eff	p
Constant	1356.9	0.000	1278.0	0.000
Gender	15.98	0.189	31.34	0.051
Age	-0.519	0.149	-0.132	0.770
Weight	*-0.858*	*0.035*	-0.366	0.468
T_2	R=0.60, p<0.001		R=0.57, p=0.001	
	Co-eff	p	Co-eff	p
Constant	34.44	0.000	34.97	0.000
Gender	-0.163	0.839	-1.175	0.142
Age	*0.074*	*0.003*	*0.067*	*0.006*
Weight	*0.073*	*0.009*	0.049	0.064
MTR	R=0.75, p<0.001		R=0.61, p<0.001	
	Co-eff	p	Co-eff	p
Constant	35.90	0.000	33.707	0.000
Gender	*-0.878*	*0.000*	0.164	0.485
Age	*-0.029*	*0.000*	*-0.032*	*0.000*
Weight	*-0.030*	*0.000*	-0.002	0.789

R overall model correlation coefficient, *Co-eff* partial regression coefficient, *p* significance level (coefficients with p<0.05 in *italics*)

predefined positioning protocol could improve scan-scan consistency [43]. Mean all-muscle summary measures provide superior reliability to individual muscle measures; an approach which would be appropriate in NMD with diffuse rather than specific muscle involvement.

Rather than assessing scan-scan reproducibility in the same session [17], a 2-week rescan interval was chosen to better simulate clinical trial conditions whilst being short enough that a true underlying physiologically-driven change in muscle MRI properties was unlikely. We did not explicitly check for factors such as recent exercise [29, 44] or diet [19], known to influence muscle T_2 and fat content respectively. Nevertheless, high reproducibility and the ability to demonstrate subtle age, weight and gender dependencies suggest that, in practice, metabolic perturbations due to typical exercise and diet regimes are small. Thus, these factors are unlikely to confound quantification of muscle pathology, an observation important for experimental trials where such factors may be hard to control.

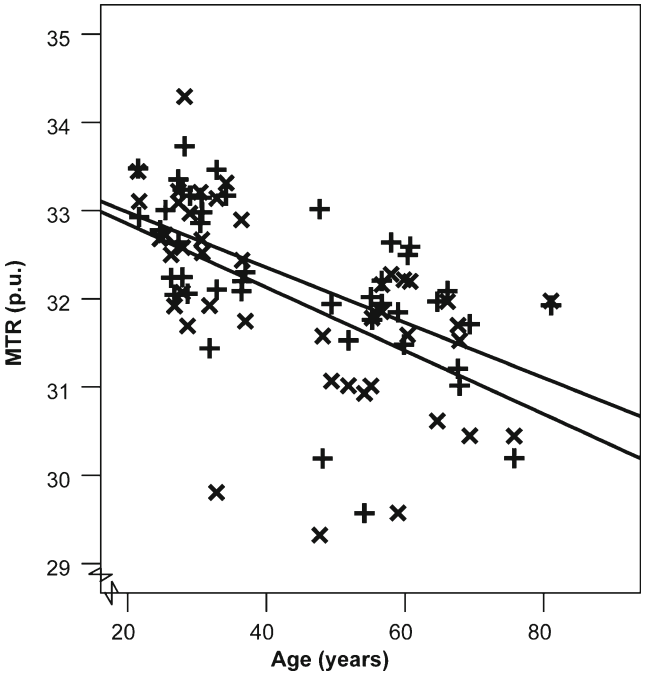

Fig. 4 Overall mean thigh (×) and calf (+) MTR is negatively correlated with subject age (p<0.001)

Age, gender and body-weight dependencies

Correlation of candidate MRI measure values with age, weight and gender is important, firstly, as such factors provide plausible surrogates for disease-related changes, usefully evidencing potential outcome measure validity. Conversely these dependencies, if severe, may confound imaging assessment of outcome by masking changes due to disease. In our healthy volunteers, consistent with age-related impaired muscle strength and neurophysiological performance [4, 45], muscle MTR reduced while T_2 and FF increased with age in both thigh and lower leg muscles. Schwenzer et al. [25] also demonstrated increases in calf-level FF and T_2 in older subjects, but not MTR. Our contrasting MTR observation may be due to acquisition condition differences, or the advantage of performing B_1 correction [35] in our study. MTR was the

Table 4 Pearson correlation coefficients between quantitative parameters in individual muscles

	T1	T2	MTR
Calf			
Dixon FF	-0.28	0.61	-0.30
T_1		-0.18	0.42
T_2			-0.47
Thigh			
Dixon FF	-0.42	0.62	-0.41
T_1		-0.21	0.48
T_2			-0.51

All correlations significant, p<0.00001

FF fat fraction, *MTR* magnetisation transfer ratio

measure most sensitive to demographic factors, the negative correlation with age being highly significant ($p<0.01$) for both overall means, and many individual muscles. The correlation remained significant in a model with T_2 and FF included as covariates, suggesting an MTR age-dependence independent of age-related muscle lipid increases, presumably reflecting myofibre quality and density changes. Future studies involving fat-suppressed or IDEAL-based measurement [46] may conclusively identify muscle-tissue water variations independent of lipid content change.

The significant associations between FF, T_2 and MTR with weight, and also between MTR and gender, in the thigh, none of which were observed in the calf-level muscle groups, presumably reflect preferential lipid accumulation in the thigh. These quantitative imaging findings are consistent with muscle lipid increases with weight [17, 18] but not gender [27] on ^{1}H-MRS. In any case, these demographically driven differences are smaller than the expected pathological changes in NMDs, and thus too small to pose a significant finding in longitudinal studies. This is in contrast with the typically wide variation present in the healthy population for neurophysiology and myometry outcome measures.

Feasibility/study limitations

To allow for straightforward application in future multi-site trials, we chose to test sequences readily implemented on standard MRI systems with unmodified software, and which can provide reasonable anatomical coverage in practical examination times. This necessarily limited the measurement sophistication, e.g. multi-echo T_2 measurement sequences allowing analysis of multiple T_2 decay components [29] did not meet the criteria of ready availability and anatomical coverage versus acquisition time. Nevertheless the sequences chosen were adequate to provide sensitive and reproducible measures of FF, T_2 and MTR relevant to muscle pathology.

A challenge in lower-limb quantitative MRI is the inherent B_1 inhomogeneity, particularly at field strengths of 3T and higher. While the dual-contrast TSE T_2-relaxometry and Dixon FF measurements used here were reasonably insensitive to this, even with B_1-correction MTR and T_1-relaxometry data were compromised in regions of maximum B_1 deviation. Despite this we were able to demonstrate strong muscle-MTR dependencies upon age, weight and gender. In this study, T_1 was the least reproducible measure, the least sensitive to demographic variations, and did not add explanatory power for these factors. We conclude that lower-limb muscle T_1 obtained using the DESPOT-1 relaxometry method may not be useful as an NMD outcome measure.

Although the T_1, T_2, FF and MTR values and healthy variations we present provide useful reference data to guide the design of future NMD MRI acquisition protocols, the specific absolute values obtained may be partially dependent on sequence design details and field strength. Quality control to ensure consistent inter-site measurement values will be an important first stage in the design of multi-centre trials incorporating MRI outcome measures. The reproducibility and sensitivity to healthy variations we obtained strongly support the potential applicability of these MRI measures to assess longitudinal disease progression.

We demonstrated the feasibility of performing a comprehensive range of MRI measurements in two anatomical levels in both lower limbs. In certain patients such measurements may not provide suitable outcome measures if pathological involvement is minimal, or already progressed to an end-state severity at these levels. Whole-body muscle MRI applications are increasingly being used for diagnostic purposes [47], and obtaining normative data from all skeletal muscle regions will be a priority in future studies. Natural history studies will identify the anatomical levels where disease progression is actively evolving in specific patient groups, allowing optimally efficient, anatomically-targeted protocols to be tailored to specific trial applications. The resulting reduction in required examination times may be crucial for harmonised use in future multiple-site trials, since long duration acquisitions may represent a problem in NMD patients with for example, cardiac or pulmonary involvement.

Conclusions

Lower-limb muscle T_2, FF and MTR measures may be obtained using readily implemented methods with sufficient reliability and sensitivity to detect subtle dependencies in health upon biological factors including muscle compartment, age, weight and gender. The observations provide strong suggestive evidence that quantitative MRI can provide practical, anatomically specific outcome measures with less potentially confounding inter-subject variation than current non-imaging measurements.

Acknowledgments The scientific guarantor of this publication is Professor Tarek A Yousry. The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article. This work was undertaken at University College London Hospitals/University College London, which received a proportion of funding from the Department of Health's National Institute for Health Research Biomedical Research Centres funding scheme. Dr Fischmann is supported by a grant from the Lorenzo-Piaggio Foundation Switzerland. This work was supported by the Medical Research Council (grant numbers G0601943, G1000347). No complex statistical methods were necessary for this paper. Institutional Review Board approval was obtained. Written informed consent was obtained from all subjects in this study. Methodology: prospective, cross sectional study, performed at one institution.

References

1. Cirak S, Arechavala-Gomeza V, Guglieri M et al (2011) Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study. Lancet 378: 595–605
2. Pareyson D, Reilly MM, Schenone A et al (2011) Ascorbic acid in Charcot-Marie-Tooth disease type 1A (CMT-TRIAAL and CMT-TRAUK): a double-blind randomised trial. Lancet Neurol 10:320–328
3. Streiner DL, Norman GR (2008) Health measurement scales: a practical guide to their development and use, 4th edn. Oxford University Press, Oxford
4. Verhamme C, van Schaik IN, Koelman JHTM, de Haan RJ, de Visser M (2009) The natural history of Charcot-Marie-Tooth type 1A in adults: a 5-year follow-up study. Brain 132:3252–3262
5. Wessig C, Koltzenburg M, Reiners K, Solymosi L, Bendszus M (2004) Muscle magnetic resonance imaging of denervation and reinnervation: correlation with electrophysiology and histology. Exp Neurol 185:254–261
6. Maillard SM, Jones R, Owens C et al (2004) Quantitative assessment of MRI T2 relaxation time of thigh muscles in juvenile dermatomyositis. Rheumatol Oxf Engl 43:603–608
7. Jurkat-Rott K, Weber M-A, Fauler M et al (2009) K+-dependent paradoxical membrane depolarization and Na+ overload, major and reversible contributors to weakness by ion channel leaks. Proc Natl Acad Sci U S A 106:4036–4041
8. Chung KW, Suh BC, Shy ME et al (2008) Different clinical and magnetic resonance imaging features between Charcot-Marie-Tooth disease type 1A and 2A. Neuromuscul Disord 18:610–618
9. Wattjes MP, Kley RA, Fischer D (2010) Neuromuscular imaging in inherited muscle diseases. Eur Radiol 20:2447–2460
10. Hiba B, Richard N, Hébert LJ et al (2012) Quantitative assessment of skeletal muscle degeneration in patients with myotonic dystrophy type 1 using MRI. J Magn Reson Imaging 35:678–685
11. Kim HK, Laor T, Horn PS, Racadio JM, Wong B, Dardzinski BJ (2010) T2 mapping in Duchenne muscular dystrophy: distribution of disease activity and correlation with clinical assessments. Radiology 255:899–908
12. Bryan WW, Reisch JS, McDonald G, Herbelin LL, Barohn RJ, Fleckenstein JL (1998) Magnetic resonance imaging of muscle in amyotrophic lateral sclerosis. Neurology 51:110–113
13. McDaniel JD, Ulmer JL, Prost RW et al (1999) Magnetization transfer imaging of skeletal muscle in autosomal recessive limb girdle muscular dystrophy. J Comput Assist Tomogr 23:609–614
14. Sinclair CDJ, Morrow JM, Miranda MA et al (2012) Skeletal muscle MRI magnetisation transfer ratio reflects clinical severity in peripheral neuropathies. J Neurol Neurosurg Psychiatry 83:29–32
15. Huang Y, Majumdar S, Genant HK et al (1994) Quantitative MR relaxometry study of muscle composition and function in Duchenne muscular dystrophy. J Magn Reson Imaging 4:59–64
16. Pfirrmann CWA, Schmid MR, Zanetti M, Jost B, Gerber C, Hodler J (2004) Assessment of fat content in supraspinatus muscle with proton MR spectroscopy in asymptomatic volunteers and patients with supraspinatus tendon lesions. Radiology 232:709–715
17. Shen W, Mao X, Wolper C et al (2008) Reproducibility of single- and multi-voxel 1H MRS measurements of intramyocellular lipid in overweight and lean subjects under conditions of controlled dietary calorie and fat intake. NMR Biomed 21:498–506
18. Schick F, Machann J, Brechtel K et al (2004) MRI of muscular fat. Magn Reson Med 47:720–727
19. Tamura Y, Watada H, Igarashi Y et al (2008) Short-term effects of dietary fat on intramyocellular lipid in sprinters and endurance runners. Metabolism 57:373–379
20. Kan HE, Scheenen TWJ, Wohlgemuth M et al (2009) Quantitative MR imaging of individual muscle involvement in facioscapulohumeral muscular dystrophy. Neuromuscul Disord 19: 357–362
21. Phoenix J, Betal D, Roberts N, Helliwell TR, Edwards RH (1996) Objective quantification of muscle and fat in human dystrophic muscle by magnetic resonance image analysis. Muscle Nerve 19: 302–310
22. Gloor M, Fasler S, Fischmann A et al (2011) Quantification of fat infiltration in oculopharyngeal muscular dystrophy: comparison of three MR imaging methods. J Magn Reson Imaging 33:203–210
23. Glover GH, Schneider E (1991) Three-point Dixon technique for true water/fat decomposition with B0 inhomogeneity correction. Magn Reson Med 18:371–383
24. Wren TAL, Bluml S, Tseng-Ong L, Gilsanz V (2008) Three-point technique of fat quantification of muscle tissue as a marker of disease progression in Duchenne muscular dystrophy: preliminary study. AJR Am J Roentgenol 190:W8–W12
25. Schwenzer NF, Martirosian P, Machann J et al (2009) Aging effects on human calf muscle properties assessed by MRI at 3 Tesla. J Magn Reson Imaging 29:1346–1354
26. Psatha M, Wu Z, Gammie FM et al (2012) A longitudinal MRI study of muscle atrophy during lower leg immobilization following ankle fracture. J Magn Reson Imaging 35:686–695
27. Ortiz-Nieto F, Johansson L, Ahlström H, Weis J (2010) Quantification of lipids in human lower limbs using yellow bone marrow as the internal reference: gender-related effects. Magn Reson Imaging 28:676–682
28. Gaeta M, Scribano E, Mileto A et al (2011) Muscle fat fraction in neuromuscular disorders: dual-echo dual-flip-angle spoiled gradient-recalled MR imaging technique for quantification–a feasibility study. Radiology 259:487–494
29. Ploutz-Snyder LL, Nyren S, Cooper TG, Potchen EJ, Meyer RA (1997) Different effects of exercise and edema on T2 relaxation in skeletal muscle. Magn Reson Med 37:676–682
30. Hatakenaka M, Ueda M, Ishigami K, Otsuka M, Masuda K (2001) Effects of aging on muscle T2 relaxation time: difference between fast- and slow-twitch muscles. Investig Radiol 36:692–698
31. De Certaines JD, Henriksen O, Spisni A, Cortsen M, Ring PB (1993) In vivo measurements of proton relaxation times in human brain, liver, and skeletal muscle: a multicenter MRI study. Magn Reson Imaging 11:841–850
32. Smith SM, Jenkinson M, Woolrich MW et al (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23:S208–S219
33. Deoni SCL, Rutt BK, Peters TM (2003) Rapid combined T1 and T2 mapping using gradient recalled acquisition in the steady state. Magn Reson Med 49:515–546
34. Stollberger R, Wach P (1996) Imaging of the active B1 field in vivo. Magn Reson Med 35:246–251
35. Sinclair CDJ, Morrow JM, Hanna MG et al (2012) Correcting radiofrequency inhomogeneity effects in skeletal muscle magnetisation transfer maps. NMR Biomed 25:262–270
36. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31:1116–1128
37. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1(8476):307–310

38. Sakurai Y, Tamura Y, Takeno K et al (2011) Association of T2 relaxation time determined by magnetic resonance imaging and intramyocellular lipid content of the soleus muscle in healthy subjects. J Diabetes Investig 2:356–358
39. Johnson MA, Polgar J, Weightman D, Appleton D (1973) Data on the distribution of fibre types in thirty-six human muscles. An autopsy study. J Neurol Sci 18:111–129
40. Meldrum D, Cahalane E, Conroy R, Fitzgerald D, Hardiman O (2007) Maximum voluntary isometric contraction: reference values and clinical application. Amyotroph Lateral Scler 8: 47–55
41. Robinson LR, Temkin NR, Fujimoto WY, Stolov WC (1991) Effect of statistical methodology on normal limits in nerve conduction studies. Muscle Nerve 14:1084–1090
42. Fischmann A, Hafner P, Fasler S et al (2012) Quantitative MRI can detect subclinical disease progression in muscular dystrophy. J Neurol 259:1648–1654
43. Fischmann A, Morrow JM, Sinclair CDJ et al (2013) Improved anatomical reproducibility in quantitative lower-limb muscle MRI. J Magn Reson Imaging. doi:10.1002/jmri.24220
44. Ababneh ZQ, Ababneh R, Maier SE et al (2008) On the correlation between T(2) and tissue diffusion coefficients in exercised muscle: quantitative measurements at 3 T within the tibialis anterior. Magma 21:273–278
45. Andreassen CS, Jakobsen J, Ringgaard S, Ejskjaer N, Andersen H (2009) Accelerated atrophy of lower leg and foot muscles–a follow-up study of long-term diabetic polyneuropathy using magnetic resonance imaging (MRI). Diabetologia 52:1182–1191
46. Janiczek RL, Gambarota G, Sinclair CDJ et al (2011) Simultaneous T_2 and lipid quantitation using IDEAL-CPMG. Magn Reson Med 66: 1293–1302
47. Kornblum C, Lutterbey G, Bogdanow M et al (2006) Distinct neuromuscular phenotypes in myotonic dystrophy types 1 and 2: a whole body high field MRI study. J Neurol 253:753–761